HAL
open science

Analyse non-paramétrique des politiques de maintenance basée sur des données des durées de vie hétérogènes

Ibrahima Dit Bouran Sibide

To cite this version:

Ibrahima Dit Bouran Sibide. Analyse non-paramétrique des politiques de maintenance basée sur des données des durées de vie hétérogènes. Autre. Université de Lorraine, 2014. Français. NNT : 2014LORR0081 . tel-01750804

HAL Id: tel-01750804

https://hal.univ-lorraine.fr/tel-01750804v1

Submitted on 29 Mar 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ DE LORRAINE

AVERTISSEMENT

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document.

D'autre part, toute contrefaçon, plagiat, reproduction illicite encourt une poursuite pénale.

Contact : ddoc-theses-contact@univ-lorraine.fr

LIENS

Code de la Propriété Intellectuelle. articles L 122. 4
Code de la Propriété Intellectuelle. articles L 335.2- L 335.10
http://www.cfcopies.com/V2/leg/leg_droi.php
http://www.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm

École doctorale IAEM Lorraine

Analyse non-paramétrique des politiques de maintenance basée sur des données des durées de vie hétérogènes

THÈSE

Presentée et soutenue publiquement le 16 Mai 2014

Pour l'obtention du

Doctorat de l'Université de Lorraine

(mention automatique)

par

Ibrahima dit Bouran SIDIBE

Composition du jury

Président:	M. Benoît IUNG	Professeur à l'Université de Lorraine
Rapporteurs:	M. Antoine GRALL	Professeur à l'Université de Troyes
	M. Daoud Ait KADI	Professeur à l'Université Laval (Canada)
Examinateur:	M. Mohamed EID	Commissariat à l'Energie Atomique
Directeurs:	M. Abdelhakim KHATAB	Maître de conférence à l'ENIM
	M. Kondo ADJALLAH	Professeur à l'ENIM

Laboratoire de Génie Industriel de Production et Maintenance LGIPM

À ceux qui m'ont cru capable.
À mes Parents et Ami(e)s.

À mon collègue de bureau,
Aux Membres du LGIPM,
à Madame la secrétaire du LGIPM,

À ceux qui m'ont aidé à me surpasser.

Sachez que les mots me manquent pour exprimer ma gratitude à votre égard. Les mots suivants ne représentent qu'une partie infime de mes émotions, sentiments de réconnaissance à votre égard. Cependant, je tiens du fond du cœur à vous écrire MERCI et MERCI pour tout. Car sans votre assistance rien de tout cela ne serait possible.

Sachez aussi qu'à vos côtés, ces années de thèse ont été magiques. Durant lesquelles, j'ai appris à élucider sans éluder, à me tromper sans nier, à trébucher sans tomber, à tomber sans rester à terre, à très bien compter et surtout sur vous.

Aux Membres du jury.

Je tiens tout d'abord à manifester chaleureusement mes remerciements aux membres du jury pour leur disponibilité.
Particulièrement aux rapporteurs, Monsieur ***A. Grall*** *professeur à l'université de Troyes et* ***D. Ait-Kadi*** *professeur à l'université Laval, dont les minutieuses relectures et multiples suggestions ont permis d'améliorer aussi bien le contenu que la forme de ce rapport.*
Je n'oublie pas également les examinateurs, Monsieur ***M. Eid*** *expert senior au commissariat à l'énergie atomique et Monsieur* ***B. Iung*** *professeur à l'université de Lorraine pour leurs critiques sur la pertinence de nos choix et sur la démarche scientifique entreprise dans cette thèse.*
Je ne pourrai terminer cette partie sans saluer et remercier chaleureusement Monsieur ***A. Khatab*** *co-directeur de la thèse pour sa disponibilité, son engagement scientifique dans ce projet et pour la confiance et le crédit qu'il m'a accordés tout au long de ce travail.*
Je tiens également à remercier Monsieur ***K.H. Adjallah*** *le directeur de cette thèse.*

Table des matières

Table des figures **ix**

Liste des tableaux **xiii**

Introduction générale **1**

I Retour d'expérience et modèle probabiliste des durées de vie d'un équipement **5**

1 Introduction et enjeu du retour d'expérience R.EX 6

2 Sources d'hétérogénéité : modèles de probabilité 8

2.1 Exploitation distribuée 8

2.2 Equipement distribué 10

3 Modèle probabiliste d'un équipement 11

3.1 Fonction de densité des durées de vie 11

3.2 Fonction de répartition 11

3.3 Fonction de fiabilité 11

3.4 Fonction de risque instantané ou taux de panne 12

3.5 Fonction de risque cumulé ou survie logarithmique 14

3.6 Fonction de fiabilité conditionnelle 14

3.7 Fonction des durées de vie résiduelle moyenne 15

4 Modèle statistique des durées de vie 16

5 Conclusion 17

II Estimation de la densité de défaillance basée sur les données **19**
1 Méthodes d'estimation paramétrique 20
1.1 Mesure de vraisemblance (ang. *Maximum Likehood Estimator* MLE) . . . 20
1.2 Méthode des Moments 25
1.3 Analyse de l'approche paramétrique 25
2 Méthode d'estimation semi-paramétrique 26
3 Méthodes d'estimation non-paramétrique 28
3.1 Les k_n plus proches voisins (Nearest Neighbors) 28
3.2 Fonction orthogonale ou par projection 29
3.3 Fonction noyau 30
3.3.1 Support Vector Machine (SVM) 30
3.3.2 Noyau de Parzen 31
4 Conclusion 32

III La méthode d'estimation du Noyau de Parzen **35**
1 Introduction : de l'histogramme à la méthode du Noyau de Parzen 36
2 Noyau de Parzen : définition et propriétés 39
2.1 Biais de l'estimateur 40
2.2 Variance de l'estimateur 41
2.3 Sensibilité de l'estimateur relativement à ses paramètres 41
2.3.1 Choix de la fonction Noyau 42
2.3.2 Choix du paramètre de lissage 42
3 Propagation de la variabilité du paramètre de lissage 44
4 Effet de bord 47
4.1 Techniques locales 47
4.2 Techniques globales 48
4.2.1 Noyaux asymétriques 48
4.2.2 Technique de transformation de Marron et Ruppert 50
4.2.3 Technique de transformation par difféomorphisme 51
5 Détermination de paramètre de lissage h 52
5.1 Méthodes basées sur *ISE* 54
5.2 Méthodes basées sur *AMISE* 56
5.3 Méthode de validation croisée biaisée 1 57
5.4 Méthode de validation croisée biaisée 2 57
5.5 Plug-in naïf 58

5.6 Du Plug-in de Park & Marron à celui Sheater & Jones 59
6 Applications numériques 60
6.1 Calcul du paramètre de lissage 60
6.2 Analyse des résultats 63
6.3 Correction de l'effet de bord 75
7 Conclusion 77

IV Analyse de la polique de maintenance de type âge pour un équipement usagé 79
1 Introduction 80
2 Modèle mathématique de la politique de maintenance 82
2.1 Modèle de probabilité 83
2.2 Modèle statistique 89
3 Exemple Numérique 90
4 Politique de maintenance de type âge avec réparation 103
4.1 Modèle mathématique avec réparation à T_1 103
4.2 Modèle avec réparation minimale avant T_1 106
5 Conclusion 108

V Analyse de la maintenance préventive retardée 111
1 Introduction 112
2 Analyse de la fiabilité et mode de défaillance 113
2.1 Cas où le paramètre de forme β est entier 114
2.2 Cas où le paramètre de forme β est réel positif 116
3 Retard de maintenance préventive 117
3.1 Effet du paramètre β sur la fonction psf 118
3.2 Effet du ratio ζ sur la fonction psf 119
4 Conclusion 123

Conclusion générale et perspectives 125

Annexe 1 : Mesures d'erreurs 131
1 Critères locaux 131
2 Critères glocaux 131

Annexe 2 : Autres méthodes de calcul du paramètre de lissage h 133
3 Vraisemblance 133

4 Minimisation dans L^1 (densité absolument intégrable) 134
5 Rule-and-thumb (règle et pouce) 135
6 Mesure basée sur la fonction de répartition 136
6.1 Paramètre de lissage constant 136
6.2 Paramètre de lissage variable 137
6.2.1 Paramètre $h(t_i)$ dépend des t_i 137
6.2.2 Paramètre $h(t)$ en fonction de t 138

Références bibliographiques **139**

Table des figures

III.1 L'histogramme classique pour l'estimation de la fonction de densité : taille des données $n = 500$ et largeur des intervalles $h = 4.5781$. 37

III.2 L'histogramme de Rosenblatt pour l'estimation de la fonction de densité : taille des données $n = 500$ et largeurs des intervalles $h = 2.4207$. 38

III.3 Dispersion après 30 simulations de calcul du paramètre de lissage h : cas d'une Weibull $W(2, 3; 15)$; taille des données 100. 63

III.4 Dispersion après 30 simulations de calcul du paramètre de lissage h : cas d'une Weibull $W(2, 3; 15)$; taille des données 500. 64

III.5 Dispersion après 30 simulations de calcul du paramètre de lissage h : cas d'une Weibull $W(2, 3; 15)$; taille des données 1000. 64

III.6 Dispersion après 30 simulations de calcul du paramètre de lissage h : cas de mélange pondéré de Weibull de même paramètre d'échelle $(35\%W(1, 5; 15) + 55\%W(2, 3; 15) + 10\%W(4, 5; 15))$; taille des données 100. 65

III.7 Dispersion après 30 simulations de calcul du paramètre de lissage h : cas de mélange pondéré de Weibull de même paramètre d'échelle $(35\%W(1, 5; 15) + 55\%W(2, 3; 15) + 10\%W(4, 5; 15))$; taille des données 500. 65

III.8 Dispersion après 30 simulations de calcul du paramètre de lissage h : cas de mélange pondéré de Weibull de même paramètre d'échelle $(35\%W(1, 5; 15) + 55\%W(2, 3; 15) + 10\%W(4, 5; 15))$; taille des données 1000. 66

III.9 Dispersion après 30 simulations de calcul du paramètre de lissage h : cas de mélange pondéré de Weibull de même paramètre de forme $(35\%W(2, 5; 15) + 55\%W(2, 5; 150) + 10\%W(2, 5; 1500))$; taille des données 100. 66

III.10 Dispersion après 30 simulations de calcul du paramètre de lissage h : cas de mélange pondéré de Weibull de même paramètre de forme $(35\%W(2,5;15) + 55\%W(2,5;150) + 10\%W(2,5;1500))$; taille des données 500. 67
III.11 Dispersion après 30 simulations de calcul du paramètre de lissage h : cas de mélange pondéré de Weibull de même paramètre de forme $(35\%W(2,5;15) + 55\%W(2,5;150) + 10\%W(2,5;1500))$; taille des données 1000. 67
III.12 Dispersion après 30 simulations de calcul du paramètre de lissage h : cas des données provenant d'un équipement composé de sous-équipements en concurrence $(Min\{W(1,5;15); W(2,3;15)\})$; taille des données 100. 68
III.13 Dispersion après 30 simulations de calcul du paramètre de lissage h : cas des données provenant d'un équipement composé de sous-équipements en concurrence $(Min\{W(1,5;15); W(2,3;15)\})$; taille des données 500. 68
III.14 Dispersion après 30 simulations de calcul du paramètre de lissage h : cas des données provenant d'un équipement composé de sous-équipements en concurrence $(Min\{W(1,5;15); W(2,3;15)\})$; taille des données 1000. 69
III.15 Dispersion après 30 simulations de calcul du paramètre de lissage h : cas basé sur les fonctions de répartition et des données provenant d'une distribution de melange de Weibull $(35\%W(2,5;15) + 55\%W(2,5;150) + 10\%W(2,5;1500))$; taille des données 100. 70
III.16 Dispersion après 30 simulations de calcul du paramètre de lissage h : cas basé sur les fonctions de répartition et des données provenant d'une distribution de melange de Weibull $(35\%W(2,5;15) + 55\%W(2,5;150) + 10\%W(2,5;1500))$; taille des données 500. 71
III.17 Dispersion après 30 simulations de calcul du paramètre de lissage h : cas basé sur les fonctions de répartition et des données provenant d'une distribution de melange de Weibull $(35\%W(2,5;15) + 55\%W(2,5;150) + 10\%W(2,5;1500))$; taille des données 1000. 72
III.18 Dispersion après 30 simulations de calcul du paramètre de lissage h : cas basé sur les fonctions de répartition et des données provenant d'un équipement composé de sous-équipements en concurrence $(Min\{W(1,5;15); W(2,3;15)\})$; taille des données 100. 73
III.19 Dispersion après 30 simulations de calcul du paramètre de lissage h : cas basé sur les fonctions de répartition et des données provenant d'un équipement composé de sous-équipements en concurrence $(Min\{W(1,5;15); W(2,3;15)\})$; taille des données 500. 74

III.20 Estimation par la transformation de Marron et Ruppert cas $\{W(0,8;0,1)\}$ taille des données 500. 75
III.21 Estimation par la transformation difféomorphisme cas $\{W(0,8;0,1)\}$ taille des données 500. 76

IV.1 Estimation de la fiabilité par la méthode du plug in de Jones-Sheather J-Sh. . 91
IV.2 Estimation de la fiabilité par la méthode du difféomorphisme ou Log 92
IV.3 Coût moyen de maintenance par unité de temps par rapport à l'âge : cas où la taille $n = 500$, l'âge initial $T_1 = 5$ et l'équipement est utilisé de l'environnement 1 vers 2. 93
IV.4 Coût moyen de maintenance par unité de temps par rapport à l'âge : cas où la taille $n = 500$, l'âge initial $T_1 = 5$ et l'équipement est utilisé de l'environnement 2 vers 1. 94
IV.5 Évolution de l'âge optimal T_0 de remplacement avec l'âge initial T_1. 95
IV.6 Évolution de coût optimal C_0 en fonction de l'âge initial T_1. 98
IV.7 Intervalle de variation de l'âge optimal à un écart type près, cas de l'estimation J-Sh (environnement 1 vers 2). 99
IV.8 Intervalle de variation de l'âge optimal à un écart type près, cas de l'estimation Log (environnement 1 vers 2). 100
IV.9 Intervalle de variation de l'âge optimal à un écart type près, cas de l'estimation J-Sh (environnement 2 vers 1). 101
IV.10 Intervalle de variation de l'âge optimal à un écart type près, cas de l'estimation Log (environnement 2 vers 1). 102

V.1 Les fonctions de survie potentielle pour les différentes valeurs de β. 119
V.2 Les fonctions de survie potentielle pour les valeurs simultanées de β et de ζ. . 120
V.3 Analyse de la fonction de survie potentielle en fonction de ζ pour la distribution $W(2,\eta,\gamma)$. 122
V.4 Abaque des iso-psf et lignes de niveau de fonction de survie potentielle. 123

Liste des tableaux

III.1 Les fonctions noyaux usuelles 39

IV.1 Définition de la distribution source des données 90
IV.2 Évaluation de l'âge de remplacement optimal T_0 (de l'environnement 1 à l'environnement 2) $T_1 = 1, 2, 3, 4, 5$ 96
IV.3 Évaluation de l'âge de remplacement optimal T_0 (de l'environnement 1 à l'environnement 2) $T_1 = 6, 7, 8, 9, 10$ 96
IV.4 Évaluation de l'âge de remplacement optimal T_0 (de l'environnement 2 à l'environnement 1) $T_1 = 1, 2, 3, 4, 5$ 97
IV.5 Évaluation de l'âge de remplacement optimal T_0 (de l'environnement 2 à l'environnement 1) $T_1 = 6, 7, 8, 9, 10$ 97

V.1 Les valeurs ζ_0 pour les distributions de Weibull de paramètre de forme β . . . 122

Introduction générale

La modélisation mathématique des politiques de maintenance a reçu une attention particulière en témoigne les travaux initiés dépuis les années 60 de Barlow, Hunter, Proschan, Rosenblatt (J.R.), Weiss et Wolman [1][2]. Actuellement, elle suscite encore d'énormes intérêts tant sur le plan industriel que de la recherche. Cet intérêt est dû d'une part à l'exigence des clients en termes de qualité, de fiabilité et de maintenabilité des équipements et d'autre part à des contraintes liées à la sécurité des hommes et au respect de l'environnement.

La littérature abonde de modèles mathématiques de politiques de maintenance. Ces modèles reposent sur une hypothèse fondamentale qui consiste à dire que les modèles probabilistes des durées de vie ou de degradation physique de l'équipement sont a priori connus. Les résultats d'une politique de maintenance sont clairement tributaires du modèle de fiabilité choisi. Un tel modèle doit être suffisamment précis pour garantir une politique de maintenance appropriée à l'équipement. Ces modèles doivent décrire fidèlement le comportement des équipements pour assurer une cohérence des résultats théoriques avec la réalité. Il reste évident que le choix d'un modèle influence significativement les résultats d'optimisation finaux et leurs interprétations.

En réalité, les paramètres des modèles de probabilité ou de dégradation choisis pour l'équipement sont inconnus et par conséquent doivent être estimés en se basant sur les données des durées vie expérimentales ou provenant de l'historique d'exploitation. Ces données résultent de l'usage de l'équipement soumis à des conditions d'exploitation qui ne sont pas nécessairement identiques. Par conséquent, les différents environnements d'exploitation induisent une hétérogénéité quant aux données collectées.

Des observations précedentes, nous distingons deux sources d'erreur principales dans l'approche classique basée sur les données dont l'une relative au choix du modèle paramétrique et l'autre à l'estimation de ses paramètres.

Notre présent travail de recherche consiste en une contribution qui permet de réduire l'erreur induite par le choix du modèle paramétrique tout en maîtrisant l'erreur résultant de la variabilité due à la méthode d'estimation. À travers ce travail, nous proposons une modélisation cohérente de la distribution des coûts directement sur la base de l'historique des données des durées vie. Celà aura comme avantage de supprimer une source d'erreur par l'usage de modèles non-paramétriques et continus. L'utilisation d'approches non-paramétriques permet en effet une prise en compte facile de l'hétérogénéité des données pour une discussion aisée de l'optimisation des politiques de maintenance.

L'ensemble de notre approche reposera ainsi directement sur les données des durées de vie collectées dans des environnements hétérogènes.

Le présent mémoire de thèse est composé de 5 chapitres en plus d'un chapitre de conclusion. Le Chapitre I introduit le concept des données des durées de vie en milieu hétérogène. Dans ce chapitre, nous définissons de manière générale l'équipement dans son environnement d'exploitation. Le chapitre réalise aussi une classification des données des durées de vie en fonction du type d'équipement et de ses conditions d'exploitation. Nous aborderons par la suite le modèle probabiliste des durées de vie d'un équipement.

Le Chapitre II présente un état de l'art des approches d'estimation de la densité de probabilité des durées de vie à partir des données collectées. Dans ce chapitre, l'accent est mis sur les hypothèses et le principe de chaque approche d'estimation. Nous réalisons une étude théorique sur l'efficacité de chaque approche d'estimation dans le cas de données des durées de vie hétérogénes. Les avantages et inconvénients de chacune des approches d'estimation seront mis en évidence. Cette étude justifiera le choix du Noyau de Parzen comme approche d'estimation exploitée dans le cadre de ces travaux de thèse.

Le Chapitre III est dédié à l'estimateur du Noyau de Parzen dans les étapes de sa construction et les hypothèses necéssaires pour sa convergence. Ce chapitre présente les méthodes de recherche des paramètres de l'approche du Noyau de Parzen et la propagation des erreurs qu'ils peuvent induire pour les calculs en aval. Ce chapitre présente également une analyse de la variabilité de chaque technique permettant la détermination des paramètres de l'approche.

Dans le Chapitre IV en exploitant les résultats du chapitre III, nous proposons un modèle mathématique d'optimisation pour une politique de maintenance. Cette politique est la politique de maintenance de type âge pour les équipements usagés. Cette politique de maintenance consiste à remplacer l'équipement à la suite d'une panne ou après une période de fonctionnement sans panne. Notre équipement est conçu pour être exploité dans différents environnements caractérisés par des degrés de sévérité distincts. Pour une telle politique de maintenance, nous proposons un modèle mathématique basé sur l'estimateur du Noyau de Parzen. Les condi-

tions d'existence d'une politique optimale de maintenance sont établies et discutées. Le critère d'optimisation est défini par les coûts de remplacement correctif et préventif. Plusieurs expérimentations sont réalisées et les résultats issus de l'estimation sont comparés à ceux provenant du modèle paramétrique réel en plus d'une analyse de leurs variabilités en fonction de la taille des données.

Dans ce même chapitre, d'autres variétés de politique de maintenance seront proposées et étudiées théoriquement. Dans un premier cas, l'équipement subit une réparation avant le début de son exploitation pour avoir un niveau de fiabilité requis. À travers cette nouvelle analyse, nous chercherons à déterminer l'existence d'un coût de réparation garantissant un fonctionnement jusqu'à la défaillance (run-to-fail) pour certains types équipements. Un second modèle est étudié et dans lequel, nous tiendrons compte du coût de la réparation de l'équipement pendant son exploitation antérieure. Cette réparation est supposée être minimale après la défaillance.

Enfin pour faciliter la mise en place de telles politiques de maintenance, nous analyserons les possibilités et les risques qu'engendre le report d'une maintenance préventive. Celà fait l'objet du Chapitre V. Le Chapitre V analysera le report de la maintenance préventive à travers une fonction risque qui fera un arbitrage entre la perte de fiabilité de l'équipement et le temps d'exploitation espéré. Pour finir le travail de recherche, une conclusion réalise une synthèse de la démarche établie. Cette conclusion est une discussion sur les limites de la modélisation proposée et définit le cadre des études futures ou perspectives que pourraient susciter ce travail de recherche.

Chapitre I

Retour d'expérience et modèle probabiliste des durées de vie d'un équipement

Ce chapitre abordera les problèmes liés à la collecte des données des durées de vie d'équipements en exploitation. Ce chapitre permettra également de mettre en relief quelques sources d'hétérogénéité des données des durées de vie. Par la suite, une modélisation des différents types des données des durées de vie sera presentée. Cette modélisation exploite les outils de probabilité et de variables aléatoires. L'ensemble de ce chapitre permet de situer le contexte et définir la problématique abordée dans ces travaux de thèse.

1 Introduction et enjeu du retour d'expérience R.EX

La complexité actuelle dans le monde économique avec ses aléas, impose aux entreprises des re-localisations (nationale ou internationale) pour un réajustement des capacités de production afin de répondre aux exigences des marchés. Dans cette logique, elles se mettent à l'abri de l'évolution rapide et incertaine de la conjoncture en répondant de manière précise à la demande dans les délais avec une maîtrise totale du coût d'exploitation des équipements. Cette réalité contraint à une logique de redéploiement des unités de production sur des sites d'exploitation distincts en terme d'organisation et de conditions d'exploitation dont les variables peuvent être fortement corrélées aux durées de vie des équipements. Les entreprises d'exploitation minières représentent un exemple concret du fait qu'elles exploitent des équipements standards sur différents continents pour différentes fonctions et pour des tâches complètement différentes. Sans omettre le fait que les équipements peuvent être successivement exploités sur plusieurs environnements au cours de leurs cycles de vie. Ces environnements impactent réellement le processus de dégradation et les durées de vie opérationnelles des équipements. Cette logique économique repose sur la maintenance distribuée induite par l'exploitation où le retour d'expérience est une agrégation de données hétérogènes relativement aux dégrés de sévérité de chacun des environnements d'exploitation (cadence, organisation, géographie, climat).

Sur le terrain, nous exploitons des équipements de grandes complexités. Ces équipements sont souvent caractérisés par une imbrication de technologies mettant en concurrence plusieurs modes et processus de défaillances dans un ou plusieurs environnements de fonctionnement. De plus, ces équipements sont régulièrement soumis à des cadences et sollicitations pouvant être variables. Ainsi, c'est sous cette disparité (environnements d'exploitation) que doivent être repensées et adaptées les politiques de maintenance pour une exploitation rationnelle des équipements. Ces politiques de maintenance doivent être élaborées afin d'optimiser les coûts liés à l'exploitation des équipements dans différents environnements. Dans nos travaux, nous supposons que l'hétérogénéité des données des durées de vie peut être due soit à une exploitation distribuée soit encore au fait que l'équipement est conçu d'une manière distribuée.

Le premier cas concerne l'exploitation distribuée. L'exploitation distribuée concerne les entreprises dont les équipements de production sont répartis sur un large espace géographique sur lesquels les variables climatiques et opérationnelles sont distinctes. En guise d'exemple, les circuits de voies de la SNCF (Société Nationale de Chemin de Fer) répartis sur l'ensemble du territoire français où l'environnement d'exploitation est différent d'une zone à une autre, dû aux variables climatiques (les étendues de variation des températures) et à la sollicitation quotidienne des composants (l'intensité du trafic).

Cette problématique d'exploitation distribuée, bien que souvent ignorée dans la modélisation de la maintenance, a fait l'objet d'étude récente, notamment dans les travaux de Kaffel [3]. L'auteur s'intéressa à la conception et au déploiement du processus de maintenance pour une efficacité opérationnelle. Alali-Alhouaij, dans ses travaux [4], se pencha également sur l'organisation et l'efficacité de la maintenance et effectua une analyse sur le dimensionnement et le positionnement géographiques des équipes de maintenance dans le cadre d'une exploitation géographiquement répartie (distribuée).

Le second cas de source d'hétérogénéité des données des durées de vie (Équipement distribué) signifie, dans le contexte de ce travail de thèse, que l'équipement admet différents modes de défaillance dû à sa construction. La durée de vie de l'équipement met en concurrence ces différents modes de défaillance relatifs aux sous-équipements de l'équipement. Cependant sous une hypothèse d'identification du mode de défaillance ou du sous-équipement défaillant, le PHM (Pronostic and Health Monitoring) de l'équipement demeure un moyen efficace pour la maîtrise de sa durée de vie. Les travaux de recherche de Ribot [5] fournissent un cadre d'application du PHM. Dans lesquels, l'auteur proposa des algorithmes cohérents de diagnostic/pronostic permettant une évaluation de la durée de vie résiduelle basée sur l'observation des composants d'équipement aéronautique.

Les points considérés ci-dessus montrent la difficulté que peuvent susciter l'optimisation de la maintenance basée sur le retour d'expérience, liée à l'hétérogénéité dans les historiques des données des durées de vie. Une telle difficulté résulte souvent du manque d'étiquetage des données relativement à l'organisation, les conditions climatiques d'exploitation ou l'assemblage technologique des équipements. Les valeurs de ces facteurs sont moins explicites, difficilement «modélisables» et souvent inaccessibles à l'analyste de la maintenance. Cependant, leurs conséquences se répercutent sur les durées de vie de l'équipement et par la suite sur la politique de maintenance à adopter. Les problèmes évoqués ci-dessus, qu'ils proviennent de l'environnement d'exploitation ou de la conception technologique de l'équipement, mettent en exergue les difficultés liées à la gestion des données des durées de vie. Ce sont ces difficultés qui expliquent et donnent les fondements et les motivations de nos travaux de recherche.

Si nous tenons pas compte de la nature de l'équipement ainsi que son environnement d'exploitation, il est évident qu'une politique ou stratégie de maintenance générique ne sera pas appropriée. En clair, une politique de maintenance issue d'une modélisation, sans considération de l'hétérogénéité des données des durées de vie, conduira à un nombre élevé d'action de maintenance corrective dans un environnement plus sévère et de maintenance préventive dans un environnement moins sévère. Une politique de maintenance générique coutera donc plus chère sur un environnement sévère (où l'équipement est susceptible d'être souvent défaillant)

que moins sévère. Dans le premier cas, les intervalles de temps entre les dates de maintenance préventive seront sur-évalués conduisant ainsi à un nombre élevé de maintenances correctives. En revanche dans le cas d'un environnement moins sévère, les intervalles entre maintenances préventives seront sous-évalués conduisant, de ce fait à un nombre élevé de maintenances préventives inutiles. Cette analyse succincte montre les limites d'une démarche classique qui repose sur la stabilité des conditions d'exploitation de l'équipement. De ce fait, une modélisation précise et rigoureuse de la loi de probabilité des durées de vie de l'équipement s'impose par l'usage de méthodes d'estimation adaptées ou par une homogénéisation des données. En effet, le but de l'analyse mathématique est l'élaboration de meilleures politiques de maintenance pour l'exploitation d'un équipement dans un environnement spécifique.

La suite du présent chapitre est organisée comme suit. La Section 2 présente une modélisation probabiliste des données des durées de vie en fonction des sources d'hétérogénéité. La Section 3 propose les modèles de probabilité caractérisant les durées de vie d'un équipement en général. Une conclusion est présentée à la section 4.

2 Sources d'hétérogénéité : modèles de probabilité

2.1 Exploitation distribuée

L'hétérogénéité dans un context d'exploitation distribuée se manifeste par une mixture de probabilités de défaillance $f_j(t)$ résultant du comportement de l'équipement en exploitation sur chaque environnement j. Lorsque les mêmes équipements sont exploités sur un nombre m d'environnements distincts (organisation, climatique, conditions de production) où le processus d'usure est susceptible d'être modifié par les variables d'exploitation, alors les données des durées de vie sont une mixture (mélange). Le modèle résultant de la mixture est un mélange pondéré de probabilités de défaillance [6]. En conséquent, le modèle est une somme de lois de défaillances $f_j(t)$ pondérées par un poids p_j répresentant le pourcentage des données des durées de vie provenant de chaque environnement j [6] :

$$f(t) = \sum_{j=1}^{m} p_j f_j(t), \tag{I.1}$$

avec

$$\sum_{j=1}^{m} p_j = 1. \tag{I.2}$$

Une illustration concrète de ce genre de données est fournie par le programme OREDA [1] (Offshore REliability DAta) développé par SINTEF Technology and Society, Departement of Safety and Reliability avec la collaboration d'une dizaine d'entreprises européennes productrices de pétrole et de gaz (Offshore & Onshore). L'objectif du programme consiste à recueillir des données des durées de vie sur les équipements de production de pétrole partout dans le monde afin de permettre le « benchmarking » et l'analyse de fiabilité, de risque ou encore de ré-conception d'un équipement pour une meilleure exploitation dans un environnement précis (Offshore & Onshore). Une autre illustration est fournie par la base de données Windstats [2] sur les équipements de production d'énergie spécifique aux composants des éoliennes de l'Allemagne et du Danemark. Ces dernières données ont permis d'extraire les informations sur le comportement des composants en Onshore pour les adapter en Offshore.

D'autres programmes d'acquision des données locaux existent en Allemagne WMEP [3] (Scientific Measurement and Evaluation Programme) & LWK [4], en Finland le VTT [5] et en Suède le Elforsk [6] sur les composants des éoliennes. Aux USA, la base de donnée de l'EPRI [7] (Electric Power Research Institute) est relative aux composants des équipements de production d'énergie en général. Une analyse sur ces historiques des données des durées de vie ne reflèterait aucun des équipements dans son environnement mais une contribution de chacun au vu de son environnement (son poids dans l'ensemble des données). Ainsi, un modèle générique aurait tendance à s'adapter au sous-ensemble le plus représenté. Une simple modélisation paramétrique de la distribution de défaillance risque d'être inappropriée lorsque l'on est en présence de plusieurs modes de défaillance. En revanche, une modélisation paramétrique robuste nécessitera une combinaison de plusieurs fonctions de défaillance avec leurs paramètres et poids respectifs. Cela conduira à des coûts élevés de calcul, c'est-à-dire, fastidieux, pénibles et souvent non convergents en plus du sur-apprentissage statistique qui peut en résulter en fonction du nombre de paramètres dans le modèle (voir le livre de Lawless [6] pour plus de détail sur les conditions de choix de modèles paramétriques).

1. www.oreda.com
2. www.windustry.org
3. www.windmonitor.iwes.fraunhofer.de
4. www.lksh.de
5. www.vtt.fi
6. www.elforsk.se
7. www.epri.com

2.2 Equipement distribué

Dans le cas d'équipement distribué, l'hétérogénéité des données est la conséquence directe de la mise en concurrence simultanée de plusieurs modes de défaillance. Un équipement distribué se présente en fiabilité comme une association en série de plusieurs sous-équipements ayant leur propre mode de défaillance. Dans cette optique, nous supposons que la défaillance d'un composant à un instant donné entraîne la défaillance de l'équipement entier. L'historique se composera entre autre de données des durées de vie provenant du composant qui tombera en premier en panne.

Contrairement aux données mélangées, l'historique des durées de vie issues des modes de défaillance en concurrence n'a pas une cohérence fréquentielle. Il demeure difficile de classer de manière précise une donnée des durées de vie observée à un des risques de défaillance en compétition. La densité de loi de panne est définie par la formule de Bayes. Lorsque m désigne le nombre de modes de défaillance en concurrence, $f_j(t)$ et $R_j(t)$ les fonctions de probabilité de défaillance et de fiabilité du sous-équipement ou du mode de défaillance j alors la fonction de densité de probabilité s'écrit :

$$f(t) = \sum_{j=1}^{m} \left(f_j(t) \prod_{k=1, k\neq j}^{m} R_k(t) \right). \tag{I.3}$$

L'hétérogénéité dans les données des durées de vie est induite par la conception de l'équipement [6]. D'autres auteurs ont proposés des modèles de risque concurrent pour deux modes de défaillance (accidentelle et usure) appelé modèle de Drenick [1] ou de Bertholon [7]. En excluant les défaillances simultanées, les durées de vie seront régies par la loi du minimum [6] entre les variables aléatoires caractérisant la loi de défaillance des sous-équipements en concurrence et peut être vue comme une mixture ponctuelle lorsque nous divisons l'Équation (I.3) par la somme des produits de fonctions de fiabilité $\sum_{j=1}^{m} \prod_{k=1, k\neq j}^{m} R_k(t)$.

Dans les sections précedentes, nous avons expliqué les origines et la nature d'hétérogénéité que peuvent avoir les données des durées de vie. Les sections suivantes donnent les fonctions continues nécessaires pour caractériser les durées de vie d'un équipement de manière probabiliste. La durée de vie de l'équipement sera supposée être une variable aléatoire continue notée X dont la loi de probabilité devrait être estimée à travers ses réalisations $\{t_i\}_{i=1,\cdots,n}$ constituant un historique des données des durées de vie.

3 Modèle probabiliste d'un équipement

Soit X la variable aléatoire qui caractérise les durées de vie d'un équipement. Cette variable aléatoire X a une loi de probabilité pouvant être définie par des fonctions continues suivantes.

3.1 Fonction de densité des durées de vie

La fonction $f(t)$ définit la densité de probabilité des durées de vie de l'équipement à un instant donné. C'est une fonction définit sur $[0, \infty[$, mesurable et de mesure égale à l'unité. La probabilité de défaillance de l'équipement entre $[t, t + dt]$ est approximativement égale au produit $f(t)dt$:

$$f(t) = \lim_{dt \to 0} \frac{P\left(t < X \leq t + dt\right)}{dt}. \tag{I.4}$$

3.2 Fonction de répartition

Cette fonction appelée aussi la fonction cumulative ou fonction de distribution définit la probabilité de défaillance de l'équipement avant l'instant t :

$$\begin{aligned} F(t) &= P\left(X \leq t\right), \\ &= \int_0^t f(u)\mathrm{d}u. \end{aligned} \tag{I.5}$$

La fonction de répartition $F(t)$ est une fonction croissante et bornée telle que :

$$\lim_{t \to 0} F(t) = 0, \tag{I.6}$$

$$\lim_{t \to \infty} F(t) = 1. \tag{I.7}$$

Indépendamment de la loi de défaillance sous-adjacente de l'équipement, la fonction de répartition établit une bijection de l'espaces des temps vers l'intervalle $]0, 1[$. En plus, sa distribution suit une loi uniforme $U\,]0, 1[$ indépendamment de la loi de la variable aléatoire X.

$$F(X) \to U\,]0, 1[\,. \tag{I.8}$$

3.3 Fonction de fiabilité

La fonction de survie de l'équipement est la probabilité complémentaire de la fonction de répartition $F(t)$. Elle définit la probabilité que l'équipement fonctionne sans défaillance (ne

tombe pas en panne avant l'instant t) jusqu'à de l'instant t.

$$\begin{aligned} R(t) &= P\left(X > t\right), \\ &= 1 - F(t), \\ &= \int_t^{\infty} f(u)\mathrm{d}u. \end{aligned} \tag{I.9}$$

Contrairement à la fonction de répartition, la fonction de survie ou de fiabilité est décroissante en fonction de la variable t. Elle vérifie les propriétés :

$$\lim_{t \to 0} R(t) = 1, \tag{I.10}$$

$$\lim_{t \to \infty} R(t) = 0. \tag{I.11}$$

La fonction de fiabilité ou de survie est une bijection entre l'espace des temps $]0, \infty[$ vers $]0, 1[$. Elle est aussi distribuée selon une loi uniforme $U\,]0, 1[$ quelle que soit la loi de X.

$$R(X) \to U\,]0, 1[\,. \tag{I.12}$$

NB : Le modèle probabiliste de l'équipement est décrit par l'une des fonctions ci-dessus : densité, répartition ou fiabilité. Les fonctions précedentes sont toutes équivalentes et permettent de caractériser complètement l'équipement en fonction de sa durée de vie.

3.4 Fonction de risque instantané ou taux de panne

La fonction $\lambda(t)\mathrm{d}t$ est la probabilité conditionnelle de défaillance de l'équipement à un instant $t+\mathrm{d}t$ sachant qu'il a survécu jusqu'à l'âge t sans défaillance. C'est une fonction positive, définie par :

$$\begin{aligned} \lim_{dt \to 0} \lambda(t)\mathrm{d}t &= P\left(t < X \leq t + dt | X > t\right), \\ &= \frac{P\left(t < X \leq t + dt; X > t\right)}{P\left(X > t\right)}, \\ &= \frac{P\left(t < X \leq t + dt\right)}{P\left(X > t\right)}, \end{aligned} \tag{I.13}$$

Des Équations (I.4) et (I.13), nous déduisons l'expression de la fonction de risque instantané ou de taux de défaillance (panne) par :

$$\lambda(t) = \frac{f(t)}{R(t)}. \tag{I.14}$$

Elle peut être constante, croissante IFR [8], décroissante DFR [9] ou une combinaison. La fonction $\lambda(t)$ est en bijection avec la fonction de fiabilité $R(t)$ ou de répartition $F(t)$ sur leurs espaces respectifs. De plus, si elle est strictement monotone alors la bijection s'étend à l'espace des temps. Par conséquent, la fonction $\lambda(t)$ permet de caractériser complètement la durée de vie de l'équipement. De l'Équation (I.14) et sachant que :

$$f(t) = -\frac{dR(t)}{dt}, \tag{I.15}$$

nous obtenons,

$$\lambda(t) = -\frac{d\log\left(R(t)\right)}{dt}. \tag{I.16}$$

L'intégration de l'Équation (I.16) permet d'établir la relation entre $\lambda(t)$ et $R(t)$ d'une part :

$$R(t) = \exp\left(-\int_0^t \lambda(u)\mathrm{d}u\right), \tag{I.17}$$

et d'autre part la variation de $R(t)$ par rapport à $\lambda(t)$,

$$\begin{aligned}\frac{dR(t)}{d\lambda(t)} &= \frac{dR(t)}{dt}\frac{dt}{d\lambda(t)},\\ &= \frac{1}{\lambda'(t)}\frac{dR(t)}{dt}.\end{aligned} \tag{I.18}$$

En plus, nous savons que

$$\frac{dR(t)}{dt} = -f(t) < 0,$$

alors le signe de $\dfrac{dR(t)}{d\lambda(t)}$ est l'opposé de celui $\lambda'(t)$. Il suffit que $\lambda'(t)$ ait le même signe par tout sur l'espace des temps t pour que la dépendance entre $(t, R(t), \lambda(t))$ soit monotone. Ceci conduit à une extension de la bijection entre $(R(t), \lambda(t))$ définie par l'Équation (I.17) jusqu'à l'espace des durées de vie. Sous cette condition, la connaissance de la fonction de risque instantanée

8. Increasing Failure Rate
9. Decreasing Failure Rate

$\lambda(t)$ permet de déterminer l'âge de l'équipement.

3.5 Fonction de risque cumulé ou survie logarithmique

La fonction de risque cumulé représente le nombre moyen de défaillances dans un intervalle de temps [8]. Elle est définie par l'intégrale de la fonction risque instantanée $\lambda(t)$.

$$\Lambda(t) = \int_0^t \lambda(u)\mathrm{d}u. \tag{I.19}$$

De l'Équation (I.16), nous obtenons :

$$\Lambda(t) = -\log\left(R(t)\right), \tag{I.20}$$

d'où son nom en anglais log-survival-function.

Contrairement à la fonction $\lambda(t)$, la fonction de risque cumulée est strictement croissante et est une bijection sur l'espace des durées de vie. Ainsi, avec la fonction de risque cumulée nous pouvons explicitement déterminer la durée de vie de l'équipement.

3.6 Fonction de fiabilité conditionnelle

La fiabilité conditionnelle quantifie la probabilité de survie de l'équipement sachant qu'il a survécu jusqu'à date t_0. Elle est définie par :

$$R_{t_0}(t) = P\left(X > t + t_0 | X > t_0\right), \tag{I.21}$$

$$= \frac{R(t+t_0)}{R(t_0)}. \tag{I.22}$$

Elle est réécrite en utilisant la fonction de risque $\lambda(t)$ par :

$$R_{t_0}(t) = \exp\left(-\int_{t_0}^{t+t_0} \lambda(u)\mathrm{d}u\right), \tag{I.23}$$

ou

$$R_{t_0}(t) = \exp\left(\Lambda(t_0) - \Lambda(t+t_0)\right). \tag{I.24}$$

De l'Équation (I.24), nous déduisons que la distribution de fiabilité conditionnelle permet aussi de caractériser l'âge de l'équipement. Autrement, lorsque l'on a $R_{t_0}(t)$ il suffit de poser $t_0 = 0$ pour déterminer la fonction de fiabilité $R(t)$ et après l'âge de l'équipement.

3.7 Fonction des durées de vie résiduelle moyenne

Elle constitue le temps moyen de fonctionnement espéré après un instant t sans défaillance. La durée moyenne restante est définie par la fonction de densité conditionnelle déduite de la fiabilité conditionnelle.

$$e(t) = E\left(X - t | X > t\right), \tag{I.25}$$

alors en utilisant la densité conditionnelle, nous obtenons :

$$\begin{aligned} e(t) &= \int_t^\infty (u-t)\,\frac{f(u)}{R(t)}\mathrm{d}u, \\ &= -\frac{\int_t^\infty (u-t)\,dR(u)}{R(t)}, \end{aligned} \tag{I.26}$$

avec une intégration par partie, le terme de droite de l'Équation (I.26) devient

$$\frac{\left[-\left(u-t\right)R(u)\right]_t^\infty + \int_t^\infty R(u)\mathrm{d}u}{R(t)}. \tag{I.27}$$

Si

$$\lim_{u\to\infty} uR(u) = 0,$$

alors la fonction des durées de vie résiduelle moyenne est donnée par :

$$e(t) = \frac{1}{R(t)}\int_t^\infty R(u)\mathrm{d}u. \tag{I.28}$$

Avec la fonction des durées de vie résiduelle, nous pouvons évaluer la fonction de fiabilité de l'équipement par intégration de sa fonction inverse $\left(\frac{1}{e(t)}\right)$ et nous obtenons la relation suivante.

$$R(t) = \frac{e(0)}{e(t)}\exp\left(-\int_0^t \frac{1}{e(u)}\mathrm{d}u\right), \tag{I.29}$$

où $e(0)$ est l'espérance mathématique de la variable aléatoire X des durées de vie de l'équipement. Cette relation entre $R(t)$ et $e(t)$ résulte de l'intégration de l'Équation (I.28). Alors, de l'Équation (I.29) est déduite de l'Équation (I.28) en inversant les deux termes de l'Équation (I.28) avant l'intégration. Le signe de la dérivée $\frac{de(t)}{dt}$ dépend de la différence entre $\lambda(t)$ et $\frac{R(t)}{\int_t^\infty R(u)du}$. Ce résultat est déduit de l'Équation (I.28). Ainsi, la bijection entre la durée de vie moyenne résiduelle et la fonction de survie $R(t)$ s'étend à l'espace des durées de vie t en fonction

du signe de la différence entre $\lambda(t)$ et $\dfrac{R(t)}{\int_t^\infty R(u)du}$. Par conséquent, avec la fonction des durées de vie moyenne résiduelle, nous pouvons déterminer explicitement l'âge t de l'équipement. La fonction des durées de vie résiduelle permet de caractériser l'équipement au même titre que les fonctions $f(t)$, $F(t)$, $R(t)$, $\lambda(t)$, $\Lambda(t)$, $R_{t_0}(t)$. Le modèle probabiliste d'un équipement repose explicitement sur la connaissance de l'une des fonctions ci-dessus lorsque la distribution de X est continue. Ces fonctions sont en réalité inconnues. Cela implique qu'on ne possède pas une expression explicite de ces fonctions. Elles doivent être estimées statistiquement sur la base des données des durées de vie indépendantes. Pour les équipements réparables, l'indépendance entre les données des durée de vie suppose une maintenance parfaite après chaque défaillance de l'équipement ou une modélisation des durées de première panne de l'équipement.

4 Modèle statistique des durées de vie

Sous l'hypothèse d'une identification précise des modes de défaillance de l'équipement, via un échantillon de données représentatives des durées de vie, l'estimation de la loi des durées de vie est bien possible par une démarche paramétrique. Mann et al., [9] ont défini une série de plusieurs types de modèles paramétriques capables de prendre en compte le mélange, la concurrence ainsi que le changement de mode de défaillance en fonction de l'âge de fonctionnement. Des exemples d'estimation basés sur les données des durées de vie collectées en exploitation sont présentés dans les travaux de Procaccia et al., [10]. Les résultats obtenus prouvent que ces modèles probabilistes sont fidèles aux processus de défaillance des équipements étudiés lorsque les durées de vie sont régies par un seul mode de défaillance. Pour les cas complexes, Les auteurs ont utilisé des mélanges et des concurrences de lois de Weibull pour modéliser la loi des durées de vie en présence de plusieurs modes de défaillance [10]. Les livres [10] , [11], [12] proposent d'excellents outils pour l'estimation paramétrique de la distribution des durées de vie. Les auteurs de ces livres fournissent une vue expérimentale de la science de l'estimation de la fiabilité basée sur les données des durées de vie. En effet, les auteurs ont établi dans leurs livres le déroulement des approches paramétriques quels que soient les modes de défaillances observés et les hypothèses sur la qualité de la maintenance de l'équipement ; c'est à dire maintenance imparfaite, remplacement ou réparation minimale.

Cependant, Lawless [13] a présenté les limites de l'hypothèse a priori de l'approche paramétrique en utilisant les données des durées de vie expérimentales des roulements à billes. Dans [13], l'auteur cherchait une modélisation de la distribution des durées de vie des roulements à billes. Lawless [13] a montré, pour les familles de distribution de probabilité Lognormale

et Weibull validant les tests d'adéquation, que les estimations des quantiles par intervalles de confiance sont distinctes avec une faible intersection. Ces intersections entre les estimations de quantile par intervalles de confiance diminuent surtout au voisinage des queues de distribution. Par conséquent, les démarches classiques basées sur les estimations paramétriques peuvent être inappropriées à cause du choix a priori de la distribution des données des durées de vie. Cette observation met en évidence l'inconvénient majeur de l'approche paramétrique. D'autres inconvénients liés à la convergence des calculs seront mis en évidence par la suite.

Il existe d'autres approches dites semi-paramétriques et non-paramétriques. Les approches semi-paramétriques se rapprochent de l'approche paramétrique dans leur hypothèse de construction. Ils nécessitent plus de types de données que les modèles paramétriques. En effet, cette dernière approche propose une modélisation intégrant le comportement physique avec le comportement probabiliste de l'équipement. En dernier nous avons les approches non-paramétriques. Les approches non-paramétriques sont exempts d'hypothèses a priori sur la loi des durées de vie de l'équipement. Elles se basent uniquement sur l'information provenant des données des durées de vie.

5 Conclusion

Ce chapitre a permis d'exposer la problématique de nos travaux de recherche. Dans ce chapitre, nous avons exposé le problème d'hétérogénéité inhérent à la collecte des données des durées de vie des équipements en exploitation. Par la suite, nous avons proposé pour chaque type d'hétérogénéité le modèle de probabilité qui régit les durées de vie des équipements. Ensuite et dans le cas général, nous avons exposé le modèle de probabilité des durées de vie d'un équipement. Ce modèle a permis de définir les relations qui existent entre les différentes fonctions de densité de probabilité, de fiabilité, de taux de panne et la fonction de distribution des durées de vie. La connaissance de l'une de ces fonctions permet en effet de déterminer les autres. Ainsi, à partir des données des durées de vie, il suffit de réaliser l'estimation de l'une de ces fonctions pour déterminer les autres. Cette estimation peut être réalisée par différentes méthodes. Le chapitre suivant présentera une revue de littérature des approches d'estimation des distributions des durées de vie. Nous avons regroupées ces approches en trois classes (paramétrique, semi-paramétrique et non-paramétrique). Dans ce même chapitre, nous décrirons les hypothèses et les difficultés d'implémentation de chacune des approches.

Chapitre II

Estimation de la densité de défaillance basée sur les données

Ce chapitre présente une revue de la littérature des méthodes statistiques d'estimation. À travers ce chapitre, nous définirons et expliquerons la construction de chaque approche d'estimation. Il s'agit des approches paramétrique, semi-paramétrique et non-paramétrique.
Les difficultés liées à la mise en oeuvre de chaque approche seront explicitées et comparées à celles des approches non-paramétriques et particulièrement à la méthode du Noyau de Parzen. Nous terminerons par une analyse critique qui justifie le choix de l'approche non-paramétrique (en occurrence la méthode du Noyau de Parzen) pour l'estimation des distributions des données des durées de vie d'équipements en exploitation.

Dans ce chapitre, nous partons d'un ensemble des données de durées de vie indépendantes afin d'évaluer la fonction de probabilité de densité des durées de vie d'un équipement. Une loi continue de probabilité qui doit fidèlement représenter la probabilité de défaillance de l'équipement étudié. Cette représentation probabiliste des durées de vie de l'équipement nécessite des outils statistiques précis pour réduire les erreurs et la variance des modèles en aval.

L'objectif de ce chapitre est de présenter un état de l'art sur les techniques et les méthodes d'estimation en mettant en relief leurs avantages et inconvénients. Ces méthodes reposent sur trois familles d'estimations à savoir paramétrique, semi-paramétrique et non-paramétrique.

1 Méthodes d'estimation paramétrique

Une approche d'estimation basée sur une méthode paramétrique nécessite la connaissance a priori de la famille de loi qui régit les données de durées de vie. Le principe d'estimation paramétrique des lois de probabilité consiste à :

1. choisir a priori un modèle $f(t;\theta)$ avec un nombre fini de paramètres θ,
2. sélectionner une mesure d'erreur à optimiser pour avoir les meilleurs paramètres (vraisemblance, moment, écart par rapport à une estimation empirique),
3. valider le modèle paramétrique choisi (test d'hypothèse ou d'adéquation).

En estimation paramétrique, les mesures usuelles sont la mesure de vraisemblance et celle des moments.

1.1 Mesure de vraisemblance (ang. *Maximum Likehood Estimator* MLE)

Cette méthode possède des propriétés asymptotiques intéressantes montrées dans les travaux de Saporta [14]. Elle consiste, à partir d'un échantillon des données de durés de vie $\{t_i\}_{i=1,\cdots,n}$ indépendant et identiquemment distribuées ($i.i.d$) à sélectionner une fonction de densité usuelle $f(t;\theta)$ comme distribution à l'origine des données et où θ représente le vecteur de paramètres de la distribution avec $\theta \in \Theta$ (où Θ est l'ensemble des valeurs possibles de θ). Le choix de la distribution a priori $f(t;\theta)$ doit répondre à trois règles selon Lawless [6] : La convenance probabiliste du modèle, la cohérence statistique entre les données et le modèle et la facilité de calcul.

L'estimation de la densité $f(t;\theta)$ passe par la détermination d'une estimation du vecteur de paramètres θ (les paramètres nécessaires pour définir la fonction de densité de probabilité

$f(t)$) à travers la mesure de la vraisemblance. La vraisemblance est une fonction de probabilité qui évalue la chance de l'évènement caractérisé par le n − uplet d'échantillon des données des durées de vie $\{t_i\}_{i=1,\cdots,n}$ pour une fonction de probabilité définie par un vecteur de paramètres θ. La fonction de vraisemblance s'écrit :

$$L\left(t_1, t_2, \cdots, t_n; \theta\right) = P\left(t_1, t_2, \cdots, t_n | \theta\right), \tag{II.1}$$

si les durées de vie $\{t_i\}_{i=1,\cdots,n}$ sont indépendants, nous obtenons alors :

$$L\left(t_1, t_2, \cdots, t_n; \theta\right) = \prod_{i=1}^{n} P\left(t_i | \theta\right). \tag{II.2}$$

Sachant que $P(t_i|\theta)$ est donné par $f(t_i|\theta)$, l'Équation (II.2) s'écrit :

$$L\left(t_1, t_2, \cdots, t_n; \theta\right) = \prod_{i=1}^{n} f\left(t_i | \theta\right), \tag{II.3}$$

La fonction de densité de probabilité $f(t;\theta)$ est une fonction de probabilité appartenant à l'une ou à un mélange (mixture) de famille de lois usuelles :

- **Loi exponentielle**

$$f(t;\theta) = \frac{1}{\eta} \exp\left(-\frac{t}{\eta}\right) \text{ avec } t > 0$$
$$\theta = \frac{1}{\eta},$$

- **Loi normale tronquée**

$$f(t;\theta) = \frac{1}{\rho\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2\right) \text{ avec } \ t > 0$$
$$\theta = (\mu, \sigma) \text{ avec } \rho = 1 - \int_{-\infty}^{0} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2\right) \mathrm{d}t,$$

- **Loi de Weibull**

$$f(t;\theta) = \frac{\beta}{\eta}\left(\frac{t-\gamma}{\eta}\right)^{\beta-1} \exp\left(-\left(\frac{t-\gamma}{\eta}\right)^{\beta}\right) \text{ avec } t > \gamma$$
$$\theta = (\beta, \eta, \gamma).$$

■ **Loi Gamma**

$$\begin{aligned} f(t;\theta) &= \frac{1}{\Gamma(\beta)}\frac{1}{\eta}\left(\frac{t-\gamma}{\eta}\right)^{\beta-1}\exp\left(\frac{t-\gamma}{\eta}\right) \text{ avec } t>\gamma \\ \theta &= (\beta,\eta,\gamma)\,, \\ \Gamma(\beta) &= \int_0^\infty x^{\beta-1}\exp(-x)\mathrm{d}x. \end{aligned}$$

La fonction $L(t_1,t_2,\cdots,t_n;\theta)$ est une variable aléatoire ayant une fonction de densité de probabilité. Par conséquent, le meilleur des vecteurs $\theta\in\Theta$ correspondrait au mode de la distribution de densité de la vraisemblance $L(t_1,t_2,\cdots,t_n;\theta)$. Alors, la vraisemblance privilégie la valeur du vecteur θ la plus probable pour la réalisation d'un échantillons données de durées de vie $\{t_i\}_{i=1,\cdots,n}$. Le meilleur des vecteurs θ doit annuler la dérivée première de la fonction $L(t_1,t_2,\cdots,t_n;\theta)$ ou de celle de son logarithme. Afin d'illustrer l'utilisation de la méthode de vraisemblance, prenons l'exemple de la distribution de Weibull à deux paramètres largement exploitée. Pour une telle distribution la mesure de vraisemblance s'écrit :

$$\begin{aligned} L(t_1,t_2,\cdots,t_n;\theta) &= \prod_{i=1}^{n}\frac{\beta}{\eta}\left(\frac{t_i}{\eta}\right)^{\beta-1}\exp\left(-\left(\frac{t_i}{\eta}\right)^{\beta}\right) \\ &= \left(\frac{\beta}{\eta}\right)^{n}\prod_{i=1}^{n}\left(\frac{t_i}{\eta}\right)^{\beta-1}\exp\left(-\left(\frac{t_i}{\eta}\right)^{\beta}\right). \end{aligned} \tag{II.4}$$

En appliquant la fonction logarithme aux membres de l'Équation (II.4), nous obtenons :

$$\log\left(L(t_1,t_2,\cdots,t_n;\theta)\right) = n\log\left(\frac{\beta}{\eta}\right)+(\beta-1)\sum_{i=1}^{n}\log\left(\frac{t_i}{\eta}\right)-\sum_{i=1}^{n}\left(\frac{t_i}{\eta}\right)^{\beta}. \tag{II.5}$$

L'objectif est de trouver un estimateur précis de $f(t)$ par l'estimation des paramètres β et η. Pour se faire, nous posons :

$$\frac{1}{\eta^{\beta}}=\alpha,$$

nous supposons que la variation de α par rapport à β est négligeable. La valeur du paramètre β qui maximise le logarithme de la mesure de vraisemblance satisfait l'équation :

$$\frac{\partial}{\partial\beta}\log\left(L(t_1,t_2,\cdots,t_n;\alpha,\beta)\right)=0,$$

nous obtenons

$$\frac{n}{\beta} + \sum_{i=1}^{n} \log(t_i) - \alpha \sum_{i=1}^{n} t_i^{\beta} \log(t_i) = 0, \tag{II.6}$$

d'où

$$\beta = n \left(\alpha \sum_{i=1}^{n} t_i^{\beta} \log(t_i) - \sum_{i=1}^{n} \log(t_i) \right)^{-1}, \tag{II.7}$$

avec

$$\frac{\partial}{\partial \beta} (t_i)^{\beta} = \frac{\partial}{\partial \beta} \exp(\beta \log(t_i)). \tag{II.8}$$

Par une dérivation du logarithme de la fonction de vraisemblance par rapport à β, nous obtenons :

$$\begin{aligned} \frac{\partial}{\partial \alpha} \log((L(t_1, t_2, \cdots, t_n; \alpha, \beta))) &= 0, \\ \frac{n}{\alpha} - \sum_{i=1}^{n} t_i^{\beta} &= 0, \\ \alpha &= n \left(\sum_{i=1}^{n} t_i^{\beta} \right)^{-1}. \end{aligned} \tag{II.9}$$

En remplaçant α par son expression dans l'Équation (II.7), nous obtenons :

$$\beta = n \left(\frac{n \sum_{i=1}^{n} t_i^{\beta} \log(t_i)}{\sum_{i=1}^{n} t_i^{\beta}} - \sum_{i=1}^{n} \log(t_i) \right)^{-1}. \tag{II.10}$$

Pour un échantillon de durées de vie $\{t_i\}_{i=1,\cdots,n}$, nous obtenons l'estimateur de β noté $\hat{\beta}$ par la résolution de l'équation :

$$G\left(\hat{\beta}\right) = n \left(\frac{n \sum_{i=1}^{n} t_i^{\hat{\beta}} \log(t_i)}{\sum_{i=1}^{n} t_i^{\hat{\beta}}} - \sum_{i=1}^{n} \log(t_i) \right)^{-1}. \tag{II.11}$$

L'estimateur $\hat{\beta}$ de β est le point fixe de la fonction numérique définie par l'Équation (II.11). La méthode du maximum de vraisemblance appliquée dans le cas de la distribution de Weibull à deux paramètres admet des propriétés pertinentes [15] dont l'existence d'une solution positive unique pour l'Équation (II.11). Dans ses travaux [15], l'auteur propose une analyse détaillée de la distribution de l'estimateur du paramètre β et montre que $\frac{\hat{\beta}}{\beta}$ est indépendant de l'échantillon et du paramètre β. Cependant, la mesure du maximum de vraisemblance peut avoir plusieurs maximums locaux ou être non bornée particulièrement lorsque le paramètre d'origine γ de la

distribution de Weibull est différent de zéro [6] [16] [17]. Ce problème engendre une difficulté quant à la convergence des méthodes d'optimisation classique telles que Newton-Raphson, sécante [18] [19]. La convergence des méthodes de Newton-Raphson et sécante dépendent du choix des valeurs initiales. De ce fait, Qiao et Tsokos [19] proposent une procédure itérative simple (SIP) à convergence géométrique pour les distributions de Weibull à deux paramètres. Cet algorithme permet de trouver les solutions de la MLE pour la loi de Weibull à deux paramètres. Une extension de cet algorithme a été proposée pour une loi de Weibull à trois paramètres [20]. Contrairement à l'algorithme de Newton-Raphson [21], l'algorithme proposé dans [19] [20] converge indépendamment des valeurs initiales. D'un autre côté les travaux d'Abbassi et al. [18], montrent les aptitudes et les difficultés inhérentes à l'estimation des paramètres par méthode du Maximum de vraisemblance (MLE) de la loi de Weibull à trois paramètres. Ces auteurs ont proposé une méta-heuristique basée sur l'algorithme du recuit simulé pour le calcul des paramètres [18]. Pour contourner les mêmes difficultés liées à la résolution du MLE, d'autres travaux se sont intéressés à des mesures plus adaptées telles que l'erreur pondérée par rapport à une estimation empirique ou continue de la fonction de répartition [16] [22] [23]. Ces difficultés deviennent plus importantes dans le cas de melange de distributions de Weibull [24] [25]. Car en plus des paramètres des distributions de Weibull, les poids p_j doivent être estimés aussi. Pour y faire face, Les auteurs [24] proposent 4 algorithmes pour le calcul des paramètres dans le cas de mélange des données de durées de vie. Parmi lesquels, nous trouvons l'algorithme alternatif, l'algorithme minimax (adaptée seulement au cas de deux lois), la méthode de régression et l'algorithme de Dempster, Laird et Rubin DLR [26]. L'algorithme DLR est aussi appelé algorithme EM pour Expectation and Maximization. L'algorithme EM alterne itérativement entre deux étapes une évaluation de l'espérance de la mesure MLE (E) et la recherche des paramètres assurant une amélioration (augmentation) de la mesure MLE (M). La performance de cet algorithme repose sur l'augmentation de la valeur MLE à chaque étape (E). Le lecteur intéressé par cet algorithme peut se reporter au livre de McLachlan et Krishnan [26]. Mclachlan and Krishnan remarquent que la convergence de l'EM induit quelques problèmes. Les problèmes qui sont traduits par une lente vitesse et un problème d'initialisation [26]. Cependant, sa performance est meilleure que celle de MLE avec une méthode d'optimisation simple. Tout comme le MLE classique, son efficacité est contrôlée par le modèle sous adjacent choisi. Avant son utilisation, il faudrait avoir une idée a priori de la famille des distributions.

1.2 Méthode des Moments

Elle est plus exploitée dans l'analyse descriptive des données statistiques. Cette méthode est basée sur un choix paramétrique antérieur et utilise une comparaison entre les expressions formelles des moments g_m et leurs expressions empiriques :

$$\begin{aligned} g_m(\theta) &= E_\theta\left(u(t)\right), \\ &= \int_0^\infty u(t) f(t,\theta) \mathrm{d}t, \end{aligned} \tag{II.12}$$

où $u(t)$ est une fonction mesurable [27]. θ est le vecteurs des paramètres de la fonction de densité f.

Le fondement théorique de cette approche résulte de la convergence asymptotique des moments empiriques vers les moments théoriques,

$$\frac{1}{n}\sum_{i=1}^{n} u(t_i) \to g_m(\theta), \tag{II.13}$$

lorsque $n \to \infty$.

La détermination des paramètres nécessite autant d'équations indépendantes que de paramètres dans la distribution choisi a priori. Le problème se résume donc à trouver un nombre suffisant d'équations pour déterminer les paramètres de la loi sous-jacente des données. Pour plus de détail, le lecteur peut se référer aux travaux de Cousineau [17]. À l'instar de la méthode de Vraisemblance, l'approche par les moments nécessite la connaissance a priori d'une loi dépendant des paramètres θ. Après une comparaison entre les moments empiriques et théoriques, nous déterminons le vecteur de paramètres θ qui vérifie toutes les égalités. Nous remarquerons que cette technique devient difficilement exploitable lorsque les expressions des moments théoriques ne sont pas linéaires en fonction des paramètres de la loi a priori.

Enfin, nous notons que les méthodes des moments ainsi que le maximum de vraisemblance sont tributaires de la complexité du système d'équations à résoudre.

1.3 Analyse de l'approche paramétrique

L'approche paramétrique est très utilisée dans la modélisation des distributions de durées de vie. Sa résolution demeure difficile pour certains types de distribution. Nous pouvons le voir à travers les travaux de la littérature où les auteurs proposent des solutions pour garantir la convergence. Parmi ces travaux, nous citons ceux developpés dans [18] [19] [20] dont les résultats reposent sur la mesure de la MLE. D'autres travaux tels que [16] [22] [23] reposent sur une

mesure d'erreur pondérée par rapport à la fonction de réparation empirique ou lisse (exemple : méthode du Noyau de Parzen). Cette dernière mesure est plus adaptée en comparaison à la mesure MLE, pour le cas de la distribution de Weibull à trois paramètres. A travers les travaux cités ci-dessus, nous nous rendons compte que l'estimation paramétrique d'une seule loi de probabilité sous-jacente suscite d'énormes difficultés (pour la distribution de Weibull). Cette difficulté augmente dans le cas de mélange de lois de durées de vie. En effet, en plus de l'identification des modes de défaillances présents dans les données (choix du modèle), s'ajoutent l'estimation du poids associé à chaque mode.

Les travaux de Krifa [28] considèrent un mélange de deux lois de distribution de Weibull afin de prédire le comportement de matériaux soumis à plusieurs processus de rupture avec un modèle à 5 paramètres. Le travail de Klein et Bertsche [29] propose un algorithme génétique comme technique de résolution en cas de mélange des données de durées de vie distribuées selon des lois de Weibull. L'algorithme génétique est utilisé dans cette analyse pour identifier le nombre de distributions sous-jacentes, les paramètres de chacune des lois et leurs poids par minimisation d'une mesure d'erreur par rapport à la fonction de répartition.

L'approche paramétrique combinée à la mesure de vraisemblance ou avec une erreur par rapport à la fonction de répartition empirique est un moyen efficace pour trouver les lois de distribution des durées de vie des équipements. Cependant, il faut noter que cette démarche repose sur une hypothèse a priori pouvant être fausse.

2 Méthode d'estimation semi-paramétrique

Dans l'approche semi-paramétrique l'hétérogénéité est prise en compte par des co-variables indépendantes de l'usure due à l'âge. Cet aspect nécessite la connaissance des facteurs (variables exogène ou explicatives) susceptibles d'impacter la distribution des durées de vie en plus des durées de vie. La fonction de risque instantanée $\lambda(t)$ est liée à la fonction de densité de probabilité $f(t)$ par :

$$f(t) = \lambda(t) \exp\left(-\int_0^t \lambda(u)\mathrm{d}u\right), \tag{II.14}$$

Les modèles semi-paramétriques doivent être identifiée pour assurer la convergence de l'estimation [30]. Les modèles usuels sont ceux de la classe à risque multiplicatif tel que :

$$\lambda(t) = \lambda_0(t)W(Z), \tag{II.15}$$

où $W(Z)$ est une fonction qui caractérise les facteurs d'influences (co-variable) et $\lambda_0(t)$ représente la fonction de risque de défaillance intrinsèque dépendant de l'âge de l'équipement. La fonction $\lambda_0(t)$ est choisie de manière à appartenir à une classe de distribution paramétrique (exponentiel, Weibull, log-normale,$\cdots$). La fonction $W(Z)$ est généralement de la forme [30] :

- log linéaire (modèle de Cox) : $W(Z) = \exp(AZ)$,
- linéaire : $W(Z) = 1 + AZ$,
- logistique : $W(Z) = \log(1 + \exp(AZ))$.

Le modèle le plus répandu est celui de Cox introduit en 1972 souvent appelé aussi le modèle à risque proportionnel (Proportional Hazard Model (PHM)) [31]. Le modéle de Cox est largement exploité en maintenance [32] [33] [34]. Le modèle de Cox se caractérise par son expression des facteurs d'influence :

$$W(Z) = \exp(AZ), \tag{II.16}$$

où A est le vecteur des facteurs d'influence défini par $A = (A_1, A_2, \cdots, A_k)$ et le coefficient A_j $\{j = 1, \cdots, k\}$ est l'effet du facteur d'influence j sur la durée de vie. La vraisemblance de Cox ne prend pas en compte $\lambda_0(t)$ et permet la détermination du vecteur A qui maximise la mesure de vraisemblance de Cox ($Lcox$). Cette mesure définit par le risque de survie conditionnelle :

$$L_{Cox}(Z) = \prod_{i=1}^{k} \frac{\exp(AZ_i)}{\sum_{j \in R_{(i)}} \exp(AZ_j)}. \tag{II.17}$$

Avec ($t_i \neq t_j$ si $i \neq j$ pour tout (i, j)), et où $R_{(i)}$ est composé des durées de vie t_j supérieure à t_i. Z_i et Z_j les valeurs prises par les facteurs d'influence à t_i et t_j. Par conséquent, les durées n'interviennent que dans la construction de $R_{(i)}$ c'est-à-dire par leur rang (statistique d'ordre). Nous notons souvent par p_i le risque de survie conditionnelle à partir de l'instant t_i.

$$p_i = \frac{\exp(AZ_i)}{\sum_{j \in R_{(i)}} (AZ_j)}. \tag{II.18}$$

Cette manière d'intégrer l'hétérogénéité à travers des variables explicatives réduit les marges d'erreur et fournit un meilleur pouvoir de prédiction même en présence d'échantillon de taille faible. Elle nécessite la gestion, en plus des données de durées de vie, des données sur les facteurs d'influence.

Dans la suite d'autres approches connues pour leurs robustesse seront exposées. Ces méthodes sont particulièrement plus appropriées au traitement des données hétérogénes contrairement aux modèles paramétriques ou semi-paramétrique dont selon Saporta [14] la quantité d'information expliquée par les estimateurs est inférieure à celle contenue dans l'échantillon de

durées de vie.

3 Méthodes d'estimation non-paramétrique

Les approches non-paramétriques sont exemptes de connaissance a priori de la loi des durées de vie. Elles ne font état d'aucune loi de distributions sous-jacentes pour les données. Elles se basent plutôt sur la structure des données dans l'échantillon pour estimer les lois caractérisant les durées de vie d'un équipement. Les approches non-paramétriques sont regroupées en deux classes à savoir, les méthodes basées sur les distances de voisinage et celles basées sur des fonctions.

3.1 Les k_n plus proches voisins (Nearest Neighbors)

Pour un échantillon des données des durées de vie $\{t_i\}_{i=1,\cdots,n}$ auquel nous associons une statistique d'ordre $\{t_{(i)}\}_{i=1,\cdots,n}$ telle que $t_{(1)} < t_{(2)} < \cdots < t_{(k)} < t_{(k+1)} < \cdots < t_{(n-1)} < t_{(n)}$. L'estimateur par la méthode des k_n plus proches voisins est défini par :

$$\hat{f}_{knn}(t) = \frac{k_n}{nd(V_n(t))}, \tag{II.19}$$

avec k_n un entier tel que $1 < k_n < n$ et $d(V_n(t))$ est la distance euclidienne entre t et son k_n^{ieme} voisin le plus proche où $t_{(1)} < t_{(2)} < \cdots < t_{(m-k_n)} < \cdots < t_{(m)} < t < t_{(m+1)} < \cdots < t_{(m+k_n)} < \cdots < t_{(n)}$. Le choix de la valeur du k_n a été étudié par Mack [35]. L'estimateur du $k_n^{ième}$ plus proche voisin de la densité converge en probabilité, en moyenne quadratique et intégrée sous les conditions suivantes (voir [36] [37]) lorsque la taille des données est grande c'est à dire $(n \to \infty)$:

$$k_n \to \infty \text{ et } \frac{k_n}{n} \to 0. \tag{II.20}$$

Si de plus la relation suivante est satisfaite alors :

$$\frac{k_n}{\log(n)} \to \infty, \tag{II.21}$$

Nous atteignons une convergence complète de la fonction d'estimation.
Cette convergence devient sûre avec la probabilité 1 quel que soit t si (voir [36])

$$\frac{k_n}{\log(\log(t))}. \tag{II.22}$$

L'estimateur des k_n plus proches voisins est asymptotiquement Normale. L'intégrale sur le support des données de $\hat{f}_{knn}(t)$ est infinie. Ceci est dû au problème de queue :

$$\lim_{t\to\infty} \hat{f}_{knn}(t) > 0, \tag{II.23}$$

$$\int_0^\infty \hat{f}_{knn}(t)\,\mathrm{d}t > 1. \tag{II.24}$$

L'estimateur des k_n^{ieme} plus proches voisins de la fonction de densité a une ressemblance avec la méthode du Noyau pour certaines valeur de k_n [38]. Cependant, l'estimateur des k_n^{ieme} plus proches voisins reste globalement moins efficace que la méthode d'estimation du Noyau de Parzen [37]. Parce qu'il a un biais plus élévé surtout vers les queues des distributions que la méthode du Noyau Parzen.

3.2 Fonction orthogonale ou par projection

L'estimateur par fonction orthogonale permet une estimation par projection pour les fonctions de densité de carrée intégrable sur un intervalle I. La fonction de densité $f(t)$ etant de carré intégrale sur un intervalle I, alors elle peut être réécrite à partir d'une base orthogonale de fonctions de carrée intégrable sur I. L'estimateur par projection de la fonction de densité $f(t)$ est :

$$\hat{f}_{ort}(t) = \sum_{i=1}^{m(n)} a_i b_i(t), \tag{II.25}$$

où a_i sont les cœfficients, $b_i(t)$ une base orthonormée de fonctions carrées intégrales sur l'intervalle I et $m(n)$ la dimension de cette base.
Les cœfficients a_i sont définies par :

$$a_i = \int_0^\infty b_j(t) dF(t). \tag{II.26}$$

Les fonctions $b_i(t)$ sont définies de manière que :

$$\int_0^\infty b_j(t) b_i(t) dt = \begin{cases} 1 \text{ si } i = j \\ 0 \text{ sinon} \end{cases} \tag{II.27}$$

Lorsque la base est définie, le problème d'estimation se résume à l'estimation des coefficients a_i à travers :

$$\hat{a}_i = \int_0^{\infty} b_i(t) dF_n(t), \tag{II.28}$$

$F_n(t)$ est l'estimateur empirique de la fonction répartition de $f(t)$. L'estimateur de la fonction de densité des durées de vie s'écrit alors :

$$\hat{f}_{ort}(t) = \sum_{i=1}^{m(n)} \hat{a}_i b_i(t). \tag{II.29}$$

Si $m(n) \to \infty$ et $\dfrac{m(n)}{n} \to 0$ lorsque $n \to \infty$ alors l'estimateur par projection converge vers la fonction densité $f(t)$ (en moyenne quadratique intégrée). Cet estimateur par projection presente l'avantage d'être moins sensible au phénomène de bord ou de Gibbs (présent dans la méthode du Noyau de Parzen). Cependant, ses estimations de la fonction de densité ne sont pas toujours positives ce qui fait que sont intégrale sur le support I est souvent différent de 1. Il faut noter que des améliorations ont été apportées par Watson [39] afin de réduire l'erreur moyenne quadratique intégrée de cet estimateur.

3.3 Fonction noyau

3.3.1 Support Vector Machine (SVM)

Le SVM ou le séparateur à vaste de marge est un outil utilisé initialement pour la discrimination/séparation supervisée. Les auteurs de [40] [41] l'ont adaptée à l'estimation de la fonction densité. Les auteurs considérent l'estimation de la fonction densité comme un problème de régression entre la fonction densité $f(t)$ (variable explicative) et sa fonction de répartition $F(t)$ (variable à expliquer).

$$Af(t) = F(t), \tag{II.30}$$

où A est un opérateur Hilbertien entre la fonction de densité $f(t)$ et sa fonction de réparation $F(t)$ (remplacée par son estimateur empirique $F_n(t)$). Les auteurs de [40] [41] montrent que ce problème de régression est équivalent à la determination de la fonction $\hat{f}_{svm}(t)$ estimateur de la fonction de densité :

$$\hat{f}_{svm}(t) = \sum_{i=1}^{n} \beta_i K\left(t, t_i\right). \tag{II.31}$$

où la fonction Noyau $K(,)$ est une densité de probabilité symetrique et les coefficients β_i obtenus par minimisation de l'équation suivante :

$$B(\beta) = \sum_{i=1}^{n} \sum_{j=1}^{n} \beta_i \beta_j K(t_i, t_j), \tag{II.32}$$

sous les contraintes suivantes

$$max|F_n(t) - \sum_{j=1}^{n} \beta_j \int_{-\infty}^{t} K\left(t_j, u\right) \mathrm{d}u|_{t=t_k} = \epsilon_k \tag{II.33}$$

et

$$\sum_{i=1}^{n} \beta_i = 1, \tag{II.34}$$

avec

$$\epsilon_k = \kappa \sigma(t_k), \tag{II.35}$$

et κ est une constante positive.

Les auteurs de [40] proposent un algorithme de résolution du problème d'optimisation donné par l'Équation (II.32). Une autre variante de cet algorithme est proposée dans [41] où les auteurs utilisent la théorie des champs moyens et la fonction de perte de Vapnik pour obtenir les coefficients β_j. Cette variante est plus subjective car elle utilise un paramètre de contrôle fixé à l'avance [41].

L'estimation par la méthode SVM est efficace d'après les travaux expérimentaux réalisés dans [40] [41]. Nous notons que les auteurs n'ont pas fait d'étude théorique sur sa convergence. Leurs études ont porté sur le développement d'algorithmes de calcul des coefficients β_j. Cependant, les auteurs montrent sur une étude expérimentale de classification basée sur l'estimation de la fonction de densité que la méthode SVM donne de meilleurs résultats que la méthode du Noyau de Parzen (sans indiqué la méthode utilisée pour le calcul du paramètre de lissage) et la méthode des k_n^{ieme} proche voisin avec $k_n = 3$.

3.3.2 Noyau de Parzen

La méthode du Noyau de Parzen est la méthode que nous utiliserons dans la suite de nos travaux pour la modélisation des données de durées de vie. Elle est formellement proche de la méthode des k_n plus proche voisin [36]. La différence est conceptuelle car pour estimer la fonction de densité à un instant t dans le cas des k_n^{ieme} voisin le plus proche, le nombre k_n de voisins proche est fixe et la distance $V(d_n(t))$ par rapport k_n^{ieme} plus proche varient. Pour

l'estimateur à Noyau de Parzen, la distance du voisinage est fixe et le nombre des données de durées de vie au voisinage ajuste l'estimation de la fonction de densité. Ces méthodes sont duales. L'efficacité majeure de la méthode du Noyau de Parzen sur la méthode des k_n plus proche voisins réside dans le calcul de l'intégrale totale des fonctions d'estimation de la densité. Cette mesure de probabilité totale est infinie pour la méthode des k_n plus voisins et égale 1 pour la méthode du Noyau de Parzen. Le fait que l'intégrale de l'estimateur de Parzen soit égale à l'unité montre qu'elle est une densité de probabilité ce qui n'est pas le cas pour les k_n^{ieme} voisin (voir Équation (II.24)).

Nous remarquons que la méthode du Noyau de Parzen a la même forme mathématique que la méthode SVM. La méthode SVM correspondrait à la méthode du Noyau de Parzen avec un paramètre de lissage variable en fonction du voisinage.
Contrairement au modèle semi-paramétrique ou paramétrique, La méthode du Noyau de Parzen n'utilise qu'un paramètre (paramètre de lissage) et une fonction Noyau. Elle a une flexibilité supérieure et une simplicité de mise en œuvre par rapport aux autres méthodes de la classe des estimateurs semi-paramétriques et paramétriques. Ces méthodes (semi-paramétriques et paramétriques) enregistrent un nombre important de paramètres dont le choix du modèle et les techniques de détermination des paramètres peuvent ne pas converger.

La méthode du Noyau de Parzen a aussi des faiblesses surtout au voisinage des bords (appelé bavure par Hominal, Deheuvels [42], effet de Gibbs par Saoudi [43] et effet de bord par Silverman, Jones, Marron, Ruppert, Hall[44][45][46][47]) pour l'estimation de fonctions définies sur un support bornés comme les fonctions de densité des durées de vie. Un effet qui n'est pas présent sur l'estimateur à fonction orthogonale. Cependant, les estimations de la densité avec les fonctions orthogonales peuvent être négatives.

4 Conclusion

Ce chapitre a permis d'exposer en détail les méthodes d'estimation des fonctions de probabilité basées sur des données des durées de vie. Ces méthodes sont regroupées en trois classes en fonctions des hypothèses nécessaires à leur implémentation. Dans chacune des classes d'estimateurs, nous avons implémenté quelques approches. Ainsi, la méthode du maximum de vraisemblance et celle des moments dans la classe des estimateurs paramétriques ont été implémentées. La méthode de Cox, appartenant à la classe des estimateurs semi-paramétriques, a été implémentée. De la classe des estimateurs non-paramétriques, les méthodes implémentées sont le $k^{ième}$ voisin plus proche, le Noyau de parzen, les fonctions orthogonales et le SVM. Les

différentes méthodes d'estimation furent également décrites et comparées en fonction de leurs aptitudes à traiter les données des durées de vie hétérogènes. Cette comparaison a permis de montrer que la méthode du Noyau de Parzen présente plusieurs avantages par rapport aux méthodes qu'elles soient de la classe des estimateurs non-paramétriques, semi-paramétrique et paramétriques. Cependant, l'inconvénient majeur de la méthode du Noyau de Parzen est la bavure lorsqu'elle est utilisée sur des supports semi-bornés ou bornés. Ces inconvénients se manifestent en particulier pour les durées de vie puisque celles-ci sont positives. Ces inconvénients ont des conséquences réelles sur les estimations qui peuvent être le changement de support et d'ordre du biais au voisinage du bord. Dans le chapitre suivant, nous décrirons la démarche d'estimation du Noyau de Parzen et nous présenterons des solutions d'amélioration pour pallier son inconvénient qu'est l'effet de bord.

Chapitre III

La méthode d'estimation du Noyau de Parzen

Ce chapitre présentera en détail la méthode d'estimation du Noyau de Parzen. Le paramètre lissage et la fonction noyau de cette méthode seront définis et leur choix sera discuté afin de garantir une estimation efficace. L'effet du choix et de la variabilité seront quantifiées pour chaque élément de l'estimateur. Ainsi, nous montrerons que, sous certaines conditions, les estimations de la fiabilité et aussi de la durée de vie moyenne d'un équipement sont peu sensibles à la variation et au choix des paramètres de la méthode du Noyau de Parzen.
Dans ce même chapitre, nous discuterons également des limites de l'approche du Noyau de Parzen sur le support semi-borné. Nous terminerons par l'analyse des techniques permettant d'apporter une amélioration sur le biais de bord induit par une telle méthode d'estimation.

1 Introduction : de l'histogramme à la méthode du Noyau de Parzen

La méthode du Noyau de Parzen est une généralisation de l'estimation par les histogrammes pour l'analyse des données aléatoires indépendantes et identiquement distribuées *(i.i.d)*. L'histogramme est une représentation fréquentielle permettant l'estimation de la densité de probabilité par classe. Les classes sont caractérisées par des intervalles à largeur constante ou variable qui subdivisent l'étendue des observations [14].

Supposons disposer d'un échantillon des données des durées de vie $\{t_1, \cdots, t_j, \cdots, t_n\}$ *(i.i.d)* de taille n. À cet échantillon, nous associons une statistique d'ordre $\{t_{(1)}, \cdots, t_{(j)}, \cdots, t_{(n)}\}$ telle que $t_{(j)} < t_{(j+1)}$ pour tout j.

La construction d'un histogramme nécessite une subdivision de l'étendue de l'échantillon définie par $Ed = t_{(n)} - t_{(1)}$ en des intervalles (ou classes) le plus souvent supposés de même largeur h tel que chaque intervalle est limité par des extrémités $[a_i, a_{i+1}[$. La densité estimée par la méthode des histogrammes est définie par le ratio du cardinal de chaque intervalle par rapport au produit $n * (a_{i+1} - a_i)$. La Figure III.1 présente un exemple d'histogramme construit sur la base des données aléatoires générées à partir d'une loi définie par le mélange de trois distributions de Weibull $W(1,5;15;0)$, $W(2,35;20;0)$ et $W(4,5;30;0)$ avec des proportions respectives de 35%, 55% et 10%. Dans ce cas les intervalles et l'estimation de la densité de probabilité sont :

$$\begin{aligned} a_0 &= 0, \\ a_i &= a_0 + i * h, \\ \hat{f}_{hist}(t) &= \frac{card([a_i, a_{i+1}[)}{nh}, \quad \text{si} \quad t \in [a_i, a_{i+1}[. \end{aligned} \tag{III.1}$$

L'analyse de la Figure III.1 montre deux inconvénients essentiels de l'estimation par l'histogramme. Le premier inconvénient est que l'histogramme associe les valeurs de densité par intervalle. C'est à dire que toutes les valeurs t d'un même intervalle auront la même densité de probabilité estimée Figure III.1. Le second inconvénient est la discontinuité de l'estimateur aux points extrémités $a_0, a_1, \cdots, a_m$ caractérisée par des sauts qui engendrent une absence de régularité. Pour remédier au premier inconvénient, Rosenblatt [14] propose l'utilisation de fenêtre mobile au lieu d'intervalles figés. L'histogramme de Rosenblatt permet de construire, autour de chaque instant, une classe de largeur h centrée sur l'instant t Figure III.2. En effet, dans cette représentation l'étendue n'est pas subdivisée en intervalles figés. Nous construisons plutôt

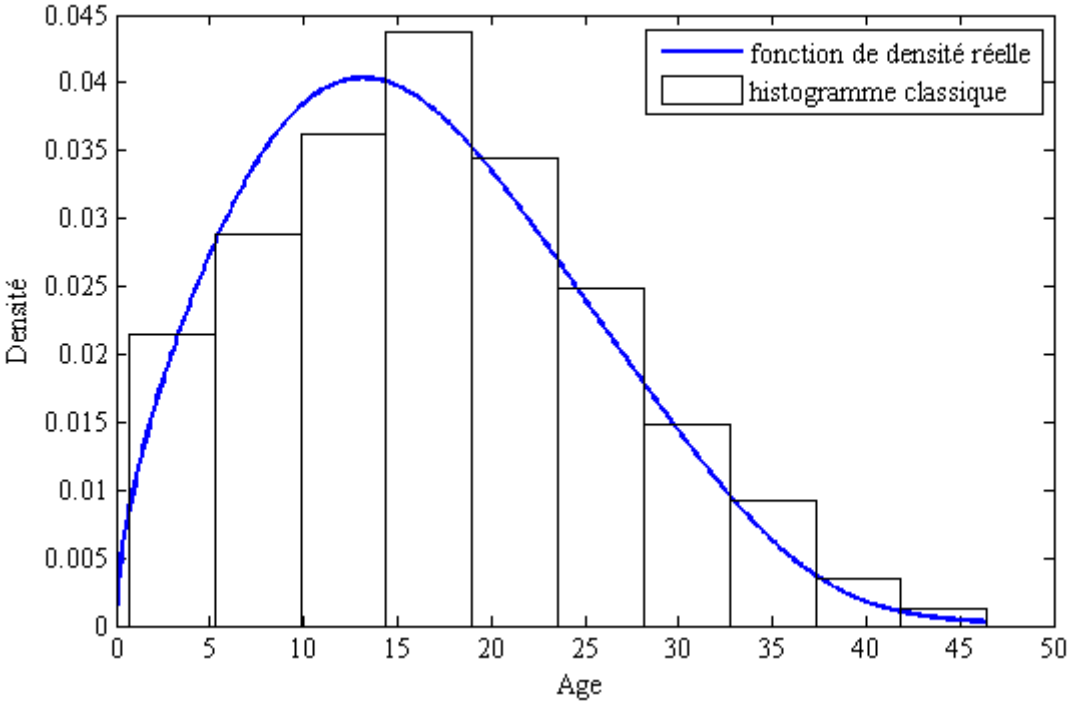

Figure III.1 – L'histogramme classique pour l'estimation de la fonction de densité : taille des données $n = 500$ et largeur des intervalles $h = 4.5781$.

un intervalle mobile centré sur chaque instant t où la densité doit être estimée. Pour chaque instant t l'intervalle mobile est défini par $I_t = \left[t - \frac{h}{2}; t + \frac{h}{2}\right]$ et caractérisant le voisinage de l'instant t. Le modèle mathématique de l'histogramme de Rosenblatt s'écrit alors :

$$\hat{f}_{ros}(t) = \frac{card(I_t)}{nh}, \tag{III.2}$$

où $Card(I_t)$ représente le nombre d'observations $\{t_i\}$ compris dans l'intervalle I_t. En utilisant la fonction indicatrice dans l'Équation (III.2), nous obtenons :

$$\hat{f}_{ros}(t) = \frac{\sum_{i=1}^{n} \mathbf{1}_{t_i \in I_t}}{nh}, \tag{III.3}$$

$$\mathbf{1}_{t_i} = \begin{cases} 1 \text{ si } t_i \in I_t, \\ 0 \text{ sinon.} \end{cases}$$

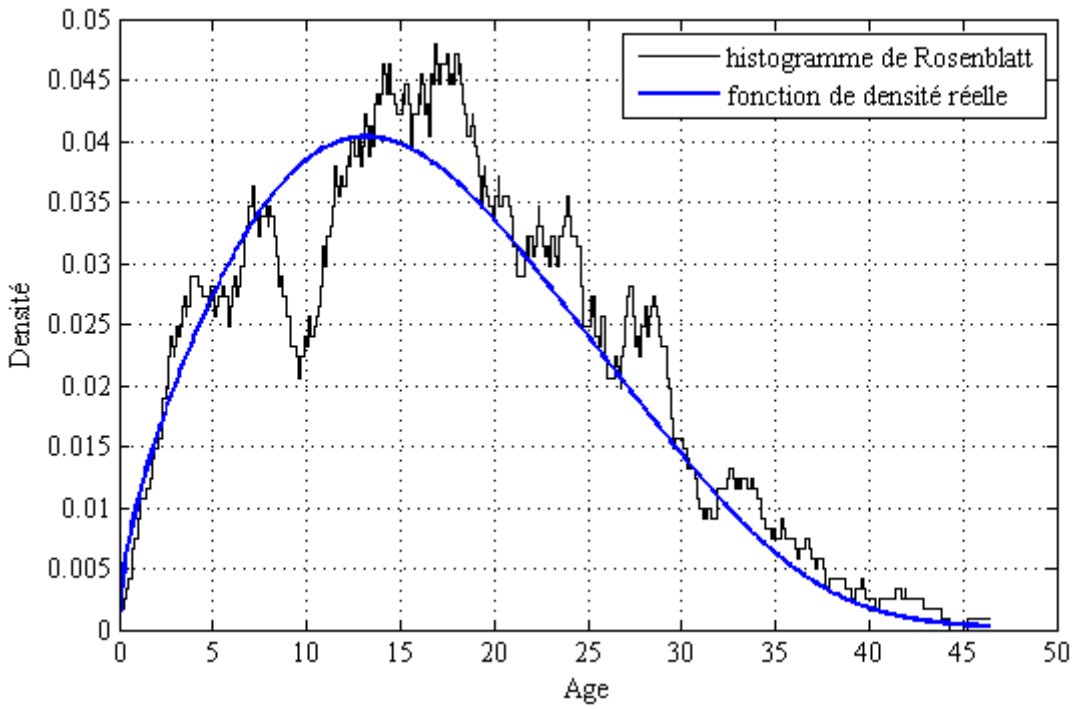

Figure III.2 – L'histogramme de Rosenblatt pour l'estimation de la fonction de densité : taille des données $n = 500$ et largeurs des intervalles $h = 2.4207$.

Si on considère la fonction $K(u)$ définie par :

$$K(u) = 1 \quad \text{si} \quad u \in \left[-\frac{1}{2}; \frac{1}{2}\right], \tag{III.4}$$

alors les Équations (III.3) et (III.2) peuvent être réecrites en remplaçant la fonction indicatrice $\mathbf{1}_{t_i}$ par la fonction $K(u)$,

$$\hat{f}_{ros}(t) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{t - t_i}{h}\right). \tag{III.5}$$

$K(u)$ représente la densité de probabilité de la loi uniforme sur $\left[-\frac{1}{2}; \frac{1}{2}\right]$, une fonction qui transmet sa discontinuité à l'estimateur de Rosenblatt (Figure III.2). Par conséquent, si nous remplaçons la fonction $K(u)$ par une fonction continue alors les estimations résulteront de la somme de fonctions continues. Ceci a été proposé dans les travaux de Parzen [48] [49].

2 Noyau de Parzen : définition et propriétés

Dans la méthode du noyau de Parzen, la fonction K est une fonction de probabilité continue vérifiant quelques propriétés mathématiques. La fonction noyau est une application de $\mathbb{R} \to \mathbb{R}$, bornée et intégrable au sens de Lebesgue [36]. Cette fonction noyau est de Geffroy si l'ensemble des points de discontinuité a une mesure nulle [36]. Elle est de Parzen-Rosenblatt si :

$$\lim_{|u| \to \infty} uK(u) = 0. \tag{III.6}$$

À l'évidence la fonction $K(u)$ vérifie :

$$\int_{-\infty}^{\infty} K(u)\mathrm{d}u = 1, \tag{III.7}$$

$$0 < \int_{-\infty}^{\infty} u^2 K(u)\mathrm{d}u < \infty. \tag{III.8}$$

Le Tableau III.1 donne quelques fonctions noyaux usuelles. Les noyaux définis dans ce tableau

Tableau III.1 – Les fonctions noyaux usuelles

Type de Noyau	Expression mathématique
Epanechnikov	$K(u) = \frac{3}{4c}\left(1 - \left(\frac{u}{c}\right)^2\right)$; $\lvert u\rvert \leq c$ où c est une constante
Normal	$K(u) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right)$
Picard	$K(u) = \frac{1}{2} \exp\left(-\lvert u\rvert\right)$
Biweight	$K(u) = \frac{15}{16}\left(1 - u^2\right)^3$, $\lvert u\rvert \leq 1$

sont symétriques et d'espérance mathématique nulle :

$$K(u) = K(-u), \tag{III.9}$$

$$\int_{-\infty}^{\infty} uK(u)\mathrm{d}u = 0. \tag{III.10}$$

De telles propriétés permettent de faciliter les dévéloppements analytiques et numériques.
Le second élément noté h de la méthode du noyau est dit paramètre de lissage. L'estimation de la fonction de densité donnée par l'Équation (III.5) est réprise avec $K(u)$:

$$\hat{f}_h(t) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{t - t_i}{h}\right). \tag{III.11}$$

La suite de cette section portera sur l'expression et l'analyse du biais et de la variance de l'estimateur du Noyau de Parzen $\hat{f}_h(t)$.

2.1 Biais de l'estimateur

L'espérance mathématique de l'estimateur du Noyau de Parzen à un instant t donné est définie par :

$$E\left(\hat{f}_h(t)\right) = \int_{-\infty}^{\infty} \hat{f}_h(t) f(u) \mathrm{d}u, \tag{III.12}$$

d'après l'Équation (III.11), l'Équation (III.12) peut s'écrire :

$$E\left(\hat{f}_h(t)\right) = \frac{1}{h} \int_{-\infty}^{\infty} K\left(\frac{t-u}{h}\right) f(u) \mathrm{d}u. \tag{III.13}$$

En effectuant le changement de variable suivant $\omega = \dfrac{t-u}{h}$, nous obtenons une autre expression de l'espérance mathématique,

$$E\left(\hat{f}_h(t)\right) = \int_{-\infty}^{\infty} K(\omega) f(t - h\omega) \mathrm{d}\omega. \tag{III.14}$$

En applicant le dévéloppement limité de Taylor à l'ordre 2 à $f(t - h\omega)$ au voisinage de t, l'Équation (III.14) devient :

$$\begin{aligned} E\left(\hat{f}_h(t)\right) &= \int_{-\infty}^{\infty} K(\omega) \left(f(t) - f'(t)\omega h + \frac{1}{2} f''(t)\omega^2 h^2 + o(h^2)\right) \mathrm{d}\omega, \\ &= f(t) \int_{-\infty}^{\infty} K(\omega) \mathrm{d}\omega - f'(t) h \int_{-\infty}^{\infty} \omega K(\omega) \mathrm{d}\omega + \frac{1}{2} f''(t) h^2 \int_{-\infty}^{\infty} \omega^2 K(\omega) \mathrm{d}\omega + o(h^2). \end{aligned} \tag{III.15}$$

Sachant que la fonction noyau $K(u)$ admet une moyenne nulle, nous obtenons alors :

$$E\left(\hat{f}_h(t)\right) = f(t) + \frac{1}{2} f''(t) h^2 \int_{-\infty}^{\infty} \omega^2 K(\omega) \mathrm{d}\omega + o(h^2). \tag{III.16}$$

Par conséquent, le biais de l'estimateur de Parzen est donné par l'approximation :

$$\begin{aligned} Bias\left(\hat{f}_h(t)\right) &= E\left(\hat{f}_h(t)\right) - f(t) \\ &= \frac{1}{2} f''(t) h^2 \int_{-\infty}^{\infty} \omega^2 K(\omega) \mathrm{d}\omega + o(h^2). \end{aligned} \tag{III.17}$$

Le terme de droite de l'Équation (III.17) est différent de zéro, ceci signifie que l'estimateur à base de Noyau de Parzen est un estimateur biaisé. A part la fonction d'impulsion de dirac (ayant $\int_{-\infty}^{\infty} \omega^2 K(\omega) d\omega = 0$), aucune autre fonction de densité de probabilité prise comme noyau ne pourrait annuler le biais défini par l'Équation (III.17).

2.2 Variance de l'estimateur

La variance de l'estimateur de Parzen s'écrit :

$$Var\left(\hat{f}_h(t)\right) = Var\left(\frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{t-t_i}{h}\right)\right), \tag{III.18}$$

Tenant compte du fait que les variables aléatoires $\{t_i\}_{i=1,\cdots,n}$ sont $i.i.d$:

$$\begin{aligned} Var\left(\hat{f}_h(t)\right) &= \frac{n}{(nh)^2} Var\left(K\left(\frac{t-t_i}{h}\right)\right), \\ &= \frac{1}{nh^2}\left(E\left[\left(K\left(\frac{t-t_i}{h}\right)\right)^2\right] - \left(E\left(K\left(\frac{t-t_i}{h}\right)\right)\right)^2\right). \end{aligned} \tag{III.19}$$

L'usage du devéloppement limité de Tailor permet d'avoir une valeur approximative des deux termes de la variance définie par l'Équation (III.19). Ainsi à $o\left(\frac{1}{nh}\right)$ près la variance est donnée approximativement par :

$$Var\left(\hat{f}_h(t)\right) = \frac{1}{nh} f(t) \int_{-\infty}^{\infty} K(\omega)^2 \mathrm{d}\omega + o(\frac{1}{nh}), \tag{III.20}$$

Les expressions approchées du biais et de la variance données respectivement par les Équations (III.17) et (III.20) montrent que le paramètre h de lissage est inversement corrélé à la variance et corrélé au biais. De plus, un paramètre de lissage élevé conduit à un biais fort et à une variance faible et inversement.

La Section 2.3 étudie en détail le problème de sensibilité de l'estimateur de Parzen par rapport à ses paramètres, la fonction noyau $K(u)$ et de la fenêtre de lissage h.

2.3 Sensibilité de l'estimateur relativement à ses paramètres

Dans cette section, nous chercherons à mesurer l'effet de chacun des paramètres de la méthode du Noyau de Parzen sur la convergence des estimations. Cette analyse de sensibilité conforte les résultats de la littérature.

2.3.1 Choix de la fonction Noyau

Des études théoriques intéressantes furent réalisées pour le choix adéquat des fonctions noyaux. Parmi lesquels, Samiuddin et El-sayyad [50] proposent une méthode basée sur la minimisation de l'erreur moyen quadratique (Mean square Error : MSE) et Cline [51] sur la minimisation de l'erreur moyenne intégrée (Mean Square Integrated Error : MISE). Pour Cline, les noyaux admissibles doivent avoir une transformée de Fourier bornée non négative et inférieur à l'unité [51]. Samiuddin et El-sayyad [50] ont fait des études comparatives entre les fonctions noyau sur le critère $MSE(K)$ en mettant en exergue que la classe de noyau d'Epanechnikov constitue les seuls noyaux admissibles. Car pour tout autre noyau, il est toujours possible d'avoir un noyau de la classe d'Epanechnikov qui est uniformément meilleur sur la base du critère $MSE(K)$. À titre d'exemple, pour les noyaux Uniforme et Normal standard nous obtenons un noyau d'Epanechnikov uniformément meilleur au sens du MSE avec respectivement $c = (10/3)^{1/5}$ et $c = (30\sqrt{\pi})^{1/5}$ (voir l'expression du noyau d'Epanecnikov dans le Tableau III.1). Nous notons que ces coefficients sont différents du coefficient du noyau optimal au sens du critère $MISE$ qui est donné par $c = \sqrt{5}$.

Une comparaison numérique des performances des noyaux sur le critère du $MISE$ fut réalisée dans [44] [52]. Cette comparaison est faite par rapport au meilleur noyau au sens du $MISE$ par Deheuvels [52]. Cette comparaison montra que la convergence de l'estimateur n'est pas affectée par le choix du noyau. Par contre, il est indispensable de retenir que certains noyaux ont des propriétés mathématiques intéressantes (dérivabilité, facilité de calcul) qui simplifient leur usage théorique et mise en œuvre lors de la programmation.

Silverman suggère l'usage du noyau Normal [44]. Ce noyau permet en effet la réduction du nombre de modes d'une distribution estimée lorsque la taille des données augmente. En plus, le noyau Normal est infiniment dérivable sur son support et s'adapte à toutes les techniques de détermination du paramètre de lissage que nous développerons dans les sections suivantes.

D'après les résultats d'ores et déjà développés dans la littérature [44], [53], [54] le choix de la fonction noyau influe peu sur les erreurs d'estimation et la convergence de l'estimateur de Parzen. Pour nos simulations dans ce rapport, nous utiliserons que le noyau Normal.

2.3.2 Choix du paramètre de lissage

Le paramètre de lissage est le second élément de la méthode d'estimation de Parzen. Il peut être déterminé de manière locale pour chaque instant t ou globale pour tout le support sur la base des informations provenant des données. Sa valeur dépend peu de la fonction noyau, d'après les travaux de simulation d'Adjabi et Cherfaoui [54] concernant quelques noyaux usuels.

Par contre comme le montre les Équations (III.17) et (III.20), il est évidemment fonction de la taille des données et joue un rôle déterminant quant à la convergence de la méthode du Noyau de Parzen [44]. Rappelons qu'un biais fort résulte d'une valeur importante de h et une variance élevée à une valeur faible de h.

Les résultats de convergence établis dans la littérature, en fonction du paramètre de lissage h reposent sur deux conditions fondamentales. La première requiert que la valeur de h tende vers 0 alors que la seconde condition permet de relier la valeur de h à celle de la taille n de l'échantillon. Cette seconde condition suppose que la valeur $nh \to \infty$. En d'autre terme, pour satisfaire la seconde condition il faut que la vitesse de convergence de h vers 0 soit inférieure à celle de $\frac{1}{n}$. De plus, nous avons les résultats de convergence suivants :

- Convergence en probabilité [55] [56] [57], presque complète pour les fonctions de densité absolument intégrable. Si
$$h \to 0 \ ; \ nh \to \infty,$$
alors
$$\int_0^\infty |\hat{f}_h(t) - f(t)| \mathrm{d}t \to 0,$$
- Convergence en moyenne quadratique (MSE) si
$$h \to 0; \ nh \to \infty,$$
alors
$$E\left(\left(\hat{f}_h(t) - f(t)\right)^2\right) \to 0.$$
- Convergence en moyenne quadratique intégrée ($MISE$), l'intégrale du MSE.
$$\int_0^\infty E\left(\left(\hat{f}_h(t) - f(t)\right)^2\right) \mathrm{d}t = \int_0^\infty Var\left(\hat{f}_h(t)\right) \mathrm{d}t + \int_0^\infty Bias^2\left(\hat{f}_h(t)\right) \mathrm{d}t,$$
ainsi pour les mêmes conditions
$$h \to 0 \ ; \ nh \to \infty,$$
alors
$$MISE \to 0,$$
- Convergence uniforme en probabilité [36]
f est une densité uniformement continue sur son support et $K(u)$ un Noyau de Parzen intégrable alors si pour $n \to \infty$ nous avons $h \to 0$ et $nh^2 \to \infty$ alors l'estimateur $\hat{f}_h$

converge en probabilité vers f.

- Convergence uniforme presque complète [36]
  nous obtenons une convergence pour les noyaux intégrables lorsque $n \to \infty$ sous la condition
  $$h \to 0, \text{ et si } \forall\ \kappa > 0 \text{ on a } \sum_{i=1}^{\infty} \exp\left(-\kappa n h^2\right) < \infty$$
  nous obtenons la même convergence avec les noyaux de Geffroy quand les conditions suivantes sont vérifiées,
  $$h \to 0, \text{ et } \frac{nh}{\log n} \to \infty.$$

En conséquence, l'estimateur du Noyau de Parzen est un estimateur asymptotiquement sans biais. Sa convergence dépend plus du paramètre de lissage h que de la fonction noyau $K(u)$.

Après une présentation de la méthode d'estimation, nous avons introduit et défini la méthode du Noyau de Parzen. Cette partie nous a permis d'identifier les paramètres de la méthode et de comprendre la criticité de chacun d'eux. Les résultats montrent que le paramètre de lissage est determinant quant à la convergence de la méthode.

La section suivante aborde la propagation et l'impact de la variation du paramètre de lissage h sur les calculs en aval.

3 Propagation de la variabilité du paramètre de lissage

Le paramètre de lissage est indispensable pour la convergence de l'estimateur du Noyau de Parzen. Il dépend de l'échantillon ainsi que de sa taille. Le paramètre de lissage constitue une variable aléatoire fonction de l'échantillon des données. Ainsi, les variations du paramètre de lissage h peuvent affecter les calculs d'optimisation en aval de l'estimation de la fonction de densité par la méthode du Noyau de Parzen. Les propositions suivantes énumèrent les conditions nécéssaires garantissant une faible propagation des erreurs (variations) du paramètre de lissage sur l'estimation les résultats en aval.

Proposition 1 *Si la fonction noyau $K(u)$ vérifie la condition suivante :*

$$\lim_{u \to \infty} uK(u) = 0, \tag{III.21}$$

alors l'effet de la variation du paramètre de lissage h sur l'estimation de la fonction de fiabilité

est borné. En plus, lorsque le noyau a une moyenne nulle :

$$\int_{-\infty}^{\infty} uK(u)\mathrm{d}u = 0, \tag{III.22}$$

alors cet effet devient nul lorsque la taille des données de durée de vie est suffisamment grande.

Preuve. Nous déduisons des Équations (I.9 de la page 12) et (III.11) l'estimateur de la fonction de fiabilité par la méthode du noyau,

$$\hat{R}_h(t) = \frac{1}{n}\sum_{i=1}^{n}\int_t^{\infty}\frac{1}{h}K\left(\frac{u-t_i}{h}\right)\mathrm{d}u, \tag{III.23}$$

sa fonction dérivée par rapport à h est

$$\frac{d\hat{R}_h(t)}{dh} = \frac{1}{n}\sum_{i=1}^{n}\int_t^{\infty}\left(-\frac{1}{h^2}K\left(\frac{u-t_i}{h}\right) - \frac{u-t_i}{h^3}K'\left(\frac{u-t_i}{h}\right)\right)\mathrm{d}u,$$

$$\frac{d\hat{R}_h(t)}{dh} = \frac{1}{n}\left(\sum_{i=1}^{n}\int_t^{\infty}-\frac{1}{h^2}K\left(\frac{u-t_i}{h}\right)\mathrm{d}u + \sum_{i=1}^{n}\frac{1}{h^2}\int_t^{\infty}-\frac{u-t_i}{h}K'\left(\frac{u-t_i}{h}\right)\mathrm{d}u\right). \tag{III.24}$$

Une intégration par partie du second terme à droite de l'égalité dans l'Équation (III.24) donne :

$$\int_t^{\infty}-\frac{u-t_i}{h}K'\left(\frac{u-t_i}{h}\right)\mathrm{d}u = \left[(u-t_i)K\left(\frac{u-t_i}{h}\right)\right]_t^{\infty} + \int_t^{\infty}K\left(\frac{u-t_i}{h}\right)\mathrm{d}u.$$

Selon la condition de l'Équation (III.21), nous déduisonson que :

$$\frac{d\hat{R}_h(t)}{dh} = \frac{1}{n}\left(\sum_{i=1}^{n}\int_t^{\infty}-\frac{1}{h^2}K\left(\frac{u-t_i}{h}\right)\mathrm{d}u + \sum_{i=1}^{n}\left[\frac{1}{h}\frac{t-t_i}{h}K\left(\frac{t-t_i}{h}\right) + \frac{1}{h^2}\int_t^{\infty}K\left(\frac{u-t_i}{h}\right)\mathrm{d}u\right]\right),$$

$$\frac{d\hat{R}_h(t)}{dh} = \frac{1}{nh}\left(\sum_{i=1}^{n}\frac{t-t_i}{h}K\left(\frac{t-t_i}{h}\right)\right).$$

De plus, si la taille n des données est suffisant grande, alors :

$$\lim_{n\to\infty}\sum_{i=1}^{n}\frac{t-t_i}{h}K\left(\frac{t-t_i}{h}\right) = \int_{-\infty}^{\infty}uK(u)\,\mathrm{d}u = 0.$$

■

Proposition 2 *Lorsque la fonction noyau $K(u)$ vérifie les conditions de l'Équations (III.21)*

et (III.22) de la Proposition (1) l'estimateur de l'espérance mathématique de la durée de vie est peu sensible à la variation du paramètre de lissage h.

Preuve. Des Équations (I.25 de la page 15) et (III.23), l'estimateur de l'espérance mathématique des durées de vie est

$$\hat{e}_h(0) = \int_0^{\infty} \hat{R}_h(u)\mathrm{d}u, \tag{III.25}$$

posons que :

$$K_s(t) = \int_t^{\infty} K(u)\mathrm{d}t,$$

alors l'estimateur de l'espérance mathématique des durées de vie par la méthode du noyau devient,

$$\hat{e}_h(0) = \int_0^{\infty} \frac{1}{n}\sum_{i=1}^{n} K_s\left(\frac{u-t_i}{h}\right)\mathrm{d}u, \tag{III.26}$$

La fonction dérivée par rapport au paramètre de lissage permet de quantifier l'impact de la variation de h sur l'espérance de la durée de vie. Avec $\dfrac{dK_s(u)}{du} = -K(u)$ sa fonction dérivée par rapport au paramètre de lissage s'écrit :

$$\frac{d}{dh}\int_0^{+\infty} \hat{R}_h(u)\,\mathrm{d}u = \int_0^{+\infty} \frac{1}{n}\sum_{i=1}^{n} \frac{d}{dh}K_s\left(\frac{u-t_i}{h}\right)\,\mathrm{d}u,$$

Il s'en suit que :

$$\begin{aligned}\frac{d}{dh}\int_0^{+\infty} \hat{R}_h(u)\,\mathrm{d}u &= \int_0^{+\infty} \frac{1}{n}\sum_{i=1}^{n} \frac{u-t_i}{h^2}K\left(\frac{u-t_i}{h}\right)\,\mathrm{d}u,\\ &= \frac{1}{nh}\sum_{i=1}^{n}\int_0^{+\infty} \frac{u-t_i}{h}K\left(\frac{u-t_i}{h}\right)\,\mathrm{d}u,\end{aligned}$$

$$\frac{d}{dh}\int_0^{+\infty} \hat{R}_h(u)\,\mathrm{d}u = \frac{1}{nh}\int_0^{+\infty}\sum_{i=1}^{n} \frac{u-t_i}{h}K\left(\frac{u-t_i}{h}\right)\,\mathrm{d}u, \tag{III.27}$$

par conséquent sous les conditions de la Proposition (1) et lorsque la taille devient importante le terme sous l'intégrale dans l'Équation (III.27) s'annule.

$$\lim_{n\to\infty} \frac{1}{nh}\sum_{i=1}^{n} \frac{u-t_i}{h}K\left(\frac{u-t_i}{h}\right) = 0.$$

■

La conséquence des Propositions (1) et (2) est que la variation du paramètre de lissage h a un faible impact sur les estimations de la fonction de la fiabilité ainsi que l'espérance mathématique.

Par les rélations établies dans la Section 3 du Chapitre 1, il est aisé d'étendre ces résultats aux autres caractéristiques de l'équipement estimées à partir des données.

4 Effet de bord

La méthode du Noyau de Parzen montre des faiblesses sur des supports bornés selon Hominal et Deheuvels [42] pour l'estimation. C'est le cas des distributions de durée de durée dont les valeurs des données ne peuvent être que positives. Pour ce type de distribution on enregistre un :

- changement de support (de $]0, \infty[$ à $]h - min\{t_i\}, \infty[$)
- masse de probabilité totale sur le support $]0, \infty[$ inférieure à l'unité,
- fort biais (ordre h au lieu de h^2 sur $]0, h[$).

Dans la littérature, nous dénombrons plusieurs techniques d'amélioration de l'effet de bord que l'on regroupe en deux classes : locale et globale.

4.1 Techniques locales

Elles consistent à adopter des modifications sur l'estimateur de Parzen sur l'intervalle de $]0, h[$. Dans ce cas, nous utilisons deux estimateurs distincts l'un sur le support $]0, h[$ et l'autre sur $]h, \infty[$. Les techniques les plus simples sont la méthode de rénormalisation [45] et la réflexion des données (voir [44]). Le biais des estimateurs obtenus par rénormalisation et réflexion sont de l'ordre $o(h)$. Ces techniques permettent plutôt de résoudre le problème de masse de probabilité et non le biais. Jones [45] montra que la technique de réflexion est meilleure que la rénormalisation sur tout le voisinage du bord sur la base de l'analyse du biais et variances des estimateurs. Cependant la technique de réflexion n'est utilisable que lorsque $f'(0) = 0$.

Ces méthodes sont de même ordre de biais que la méthode de Parzen classique. Plusieurs contributions sur les deux classes existent : il s'agit des speudo-data de Cowling et Hall [47] pour la réflexion et la généralisation de la rénormalisation par la méthode de Jackknife (canif, couteau de poche ou de l'armée Suisse).

La généralisation par Jackknife [45] [58] consiste à utiliser une combinaison (avec des coefficients locaux dépendant de la distance par rapport au bord) d'estimateurs à noyau au voisinage du bord afin de garantir le même ordre de biais $o(h^2)$ sur tout le domaine de définition (au voisinage ainsi qu'à l'intérieur du domaine). L'ensemble des modèles usuels de généralisation de la méthode du noyau par Jackknife sont décrits dans [45]. Jones [45] expose pour chacun des modèles l'expression de leurs biais et variances approximatifs et souligne le fait que la cor-

rection du biais par renormalisation peut conduire à des estimations de fonction de densité négatives [58]. Jones [58] proposa un estimateur dont le biais est de l'ordre $o(h^2)$ et basé sur la renormalisation exponentielle. Les estimations de cette dernière renormalisation sont toujours positives. Cependant, la performance de ces techniques est limitée, nous pouvons se rendre compte à travers les simulations réalisées par Jones [58] sur les distributions Gamma, au sens du critère $MISE$.

L'amélioration basée sur les pseudo-data de Cowling et Hall [47] consiste à faire une réflexion plus formelle. Contrairement à la réflexion classique, Cowling et Hall [47] proposèrent une reproduction des données basée sur un ordre statistique associé à l'échantillon des données. Les auteurs ont proposé une estimation polynomiale des quantiles de $|-i/n|$ à générer (les i sont choisis de manière que $t_i < h$). Cette technique implique certes la résolution de système d'équations pour le calcul des nouvelles données mais s'adapte au cas où $f'(0) \neq 0$. Elle a un biais inférieur à celui de la réflexion simple selon Cowling et Hall [47]. Cette technique donne toujours des estimations positives.

Les techniques proposées dans ce paragraphe améliorent le biais à travers deux techniques de base la réflexion des données et la renormalisation. Elles restent locales avec une efficacité supérieure à la méthode de Parzen pour les données définies sur un support borné. Cependant, ces techniques peuvent conduire à des divisions par zéro et de masse de probabilité totale distincte de 1.

4.2 Techniques globales

Elles vont de l'usage de noyaux asymétriques aux transformations des données pour réduire les effets de bord. Nous dénombrons les techniques basées sur les noyaux Gamma & Beta de Chen [59], [60] et les méthodes de transformation des données de Marron et Ruppert [46] et de Saoudi et al., [43].

4.2.1 Noyaux asymétriques

Selon Chen [59] [60], une partie du biais de bord est expliquée par l'usage de noyau symétrique tels que le noyau Normal, d'Epanechnikov et la faible vitesse de convergence de la suite $n^{-1/5}$ vers zéro. Chen utilise des noyaux asymétriques tels que le noyau Beta ou Gamma. Le noyau Beta lorsque le support de la fonction de distribution est $[0, 1]$ et le noyau Gamma pour les fonctions à support semi-borné $[0, \infty[$. À titre d'exemple Adjabi et Lagha [61] ont réalisé une analyse non-paramétrique des durées dans les files d'attente basée sur le noyau Gamma.

La consistance du choix du noyau Beta dans l'estimateur de Parzen a été prouvée par Bouezmarni et Rolin [62]. Les auteurs étudiérent les erreurs absolues et montrèrent que cet estimateur est asymptotiquement sans biais pour les fonctions de densité de probabilité $f(t)$ deux fois dérivable. Bouezmarni et Rolin [62] prouve que l'estimation avec de tel noyau est sans biais lorsque $f(t)$ représente la distribution uniforme sur $[0,1]$. Le travail Chen dans [60] sur le noyau Beta sont identiques à celui du noyau Gamma dans [59]. Cheng prouva que cette méthode ne subit pas de variations de l'ordre du biais au voisinage des bords. L'estimateur de Cheng [59] sur le support $[0,\infty[$ est défini par :

$$\hat{f}_h(t_i) = \frac{1}{n}\sum_{j=1}^{n} K_{t_i/h+1,h}(t_j), \tag{III.28}$$

où

$$K_{t_i/h+1,h}(t_j) = \frac{t_j^{t_i/h}\exp\left(\frac{-t_j}{h}\right)}{h^{t_i/h+1}\Gamma\left(\frac{t_i}{h}+1\right)}. \tag{III.29}$$

L'expression de son biais est,

$$Biais(\hat{f}_h(t_i)) = h\left(f'(t_i) + \frac{1}{2}f"(t_i)\right) + o(h). \tag{III.30}$$

Le biais de cet estimateur est de l'ordre $o(h)$ et contient $f'(t_i)$ et $f"(t_i)$. En remarquant que t_i représente le mode de la distribution et non sa moyenne. Cheng [59] proposa une autre fonction Gamma $K_{\phi h(t),h}$ définie par morceau telle que :

$$\phi h(t_i) = \begin{cases} \frac{t_i}{h} \text{ si } t_i \geq 2h, \\ \left(\frac{t_i}{2h}\right)^2 \text{ si } t_i \in [0, 2h[. \end{cases} \tag{III.31}$$

L'estimateur du noyau de l'Équation (III.28) devient,

$$\hat{f}(t_i) = \frac{1}{n}\sum_{j=1}^{n} K_{\phi h(t_i),h}(t_j), \tag{III.32}$$

avec un biais approximatif de

$$Biais(\hat{f}_h(t_i)) = \begin{cases} \frac{1}{2}t_i f"(t_i)h + o(h), \text{ si } t_i \geq 2h \\ (1-t_i)\frac{\phi h(t_i)-t_i/h}{1+h\phi h(t_i)-t_i}f'(t_i) + o(h), \text{ si } t_i \in [0, 2h[. \end{cases} \tag{III.33}$$

Cheng dans [59] montra que le paramètre de lissage est à constante multiplicative près égale

$n^{-2/5}$ (celui des noyaux symétriques est de l'ordre $n^{-1/5}$). Nous remarquons que $n^{-2/5}$ converge plus rapidement vers 0 que $n^{-1/5}$. Les estimations par des noyaux asymétriques restent positives [59], [60].

4.2.2 Technique de transformation de Marron et Ruppert

La technique de Transformation des données fut utilisée par Wand et Devroye [63] pour l'estimation des fonctions de densité absolument intégrable et par Silverman [44] pour l'estimation des fonctions de densité carré intégrable. Dans cette technique, nous transformons les données par une fonction pour réduire l'effet de bord par la variation de la fonction de transformation. Marron et Ruppert [46] proposèrent des transformations strictement croissantes qui s'annulent à $t = 0$ et dont les dérivées première et seconde existent. La transformation de Marron et Ruppert par rapport aux données semi-bornées est une intégrale d'exponentielle définie par :

$$g(t) = \int_0^t c_1 + c_2 \exp\left(-c_3 u\right) \mathrm{d}u, \tag{III.34}$$

où $c_3 = \dfrac{1}{median\{t_i\}}$, et c_1 et c_2 solutions du système d'équations :

$$c_1 + c_2 = b_0 \tag{III.35}$$

$$-c_3 c_2 = b_1 \tag{III.36}$$

Les paramètres b_0 et b_1 sont issus d'un algorithme nommé D et proposé par Marron et Ruppert [46]. L'estimateur de Parzen devient (avec une réflexion seulement au point $t = 0$) :

$$\hat{f}_{gh}(t) = \frac{|g'(t)|}{nh} \sum_{i=1}^{n} K\left(\frac{g(t) - g(t_i)}{h}\right). \tag{III.37}$$

L'algorithme D selon Ruppert et Marron [46] nécessite deux paramètres de lissages indépendants. Le premier paramètre de lissage noté h_1 n'accepte pas toute les valeurs. Certaines valeurs de h_1 ne permettent pas de trouver une fenêtre de lissage pour les données transformée. A cet effet nous avons pris une valeur de h_1 ($h_1 = mediane\{t_i\}_{i=1\cdots n}$) pour assurer la convergence de l'algorithme du plug-in. D'un autre côté, la méthode de Marron et Ruppert n'est pas efficace pour les mélanges des données [46] une correction local est apportée par les auteurs dans un algorithme bis appelé $D - LB$ (de la même famille que la rénormalisation). Karunamuni et Albert [64] étendent la réflexion au-delà du point $t = 0$. Ces auteurs ont réalisé un mixage entre la réflexion sur $[0, h]$ et la transformation polynomiale des données pour construire un

estimateur dont le biais est de l'ordre de $o(h^2)$ sur tout le support des données. La fonction de transformation de Karunamuni et Albert est un polynôme de degré 3.

4.2.3 Technique de transformation par difféomorphisme

Les transformations par difféomorphisme ont été étudiées par Saoudi et al., [43] afin de résoudre le phénomène de Gibbs (appellation de l'effet de bord en traitement de signal) dans le cadre d'estimation de fonction de densité des signaux. L'estimateur de Saoudi et al., est identique à celui proposé par Marron et Ruppert.

Saoudi et al., [43] proposèrent des fonctions de transformation $g(u)$ difféomorphisme vers $\mathbb{R}$.

Définition nous disons que $g(u)$ est un difféomorphisme si seulement si

- $g(u)$ est une fonction bijective
- $g(u)$ est différentiable sur $]0, \infty[$
- sa réciproque est différentiable sur $\mathbb{R}$

Saoudi et al., [43] montrèrent qu'il est impossible de trouver une fonction de transformation permettant de supprimer le biais de l'estimateur.

Lorsque la transformation est une diffeomorphisme, $f(t)$ de carrée intégrable et $|g|of(t)$ intégrable alors l'estimateur par transformation est asymptotiquement sans biais, consistant, convergent. Contrairement à Ruppert et Marron [46] ou Karunamuni et Albert [64] les fonction de transformations des données sont des difféomorphismes. Ces fonctions sont des applications bijectives de $]0, \infty[$ vers $\mathbb{R}$ dont leurs dérivées tendent vers infini au voisinage du bord et admettent une fonction réciproque. Une condition nécessaire pour assurer une réduction du biais au voisinage des bords. Un exemple de difféomorphisme est la transformation logarithmique utilisé par Silverman [44] et Wand et Devroye [63] pour les données positives. Le paramètre de lissage est obtenu par minimisation numérique d'erreurs quadratique.

Dans la section suivante, nous nous intéresserons aux techniques de détermination du paramètre h relativement à l'optimisation des mesures d'erreur (voir annexe). Nous notons qu'il existe plusieurs mesures d'erreur pour la recherche du paramètre de lissage adéquate telles que celles définies pour les fonctions de densité absolument ou de carré intégrable (respectivement L^q avec $q = 1, 2$) ou la mesure de vraisemblance et de KullBack-Leiber [44][65] [66].

Sous des conditions sur le paramètre de lissage h, la convergence des écarts (estimateur de Parzen et la densité réelle) vers 0 dans L^q fut établie. A cet effet, nous pourrions se référer aux travaux de Parzen [48], [49], Silverman [44], Bosq [36], Wang et Hall [67] ou de Devroye [55]. De plus, Silverman [44] a établi une équivalence entre la distance de Kullback-Leiber et une

estimation de l'*ISE* (erreur quadratique intégrée).

5 Détermination de paramètre de lissage h

Nous rappelons que le choix du paramètre de lissage contrairement à celui du noyau est indispensable pour assurer la convergence de la méthode du Noyau de Parzen. Le paramètre de lissage ainsi que la convergence de l'estimation dépendent de la taille de l'échantillon. Il peut être obtenu par maximisation de la mesure de vraisemblance (voir annexe), par minimisation d'erreurs absolues (fonction de densité absolument intégrable ou dans L^1) ou quadratiques (fonction de densité carrée intégrable ou dans L^2).

L'efficacité du lissage dépend explicitement de la valeur de ce paramètre. Une analyse rapide des fonctions variance et biais de l'estimateur de Parzen montre qu'un paramètre de lissage petit entraine une courbe rugueuse (effet de sur-apprentissage) et une large valeur induit une courbe trop lisse (une perte d'information). Alors, le choix de ce paramètre doit se faire de manière judicieuse et en réalisant un arbitrage entre la perte d'information (biais) et le sur-apprentissage (variance). Pour l'usage de la méthode du noyau de Parzen dans le cadre de nos travaux, nous utiliserons la minimisation d'erreurs quadratiques au depend d'erreurs absolues. Il est plus facile de réaliser un algorithme de minimisation des erreurs quadratiques qu'absolues. Pour le cadre de cette thèse, nous utiliserons les techniques de calcul du paramètre de lissage basée sur les erreurs quadratiques afin de réaliser la modélisation de la densité de probabilité des durée de vie des équipements. L'analyse d'erreurs quadratiques est largement exposée dans [36][44]. Dans cette thèse, nous nous intéresserons plutôt à des paramètres issus des erreurs quadratiques principalement composé de l'*ISE* (Integrated Square Error), *MISE* (Mean Integrated Square Error) pour une minimisation globale et *MSE* (Mean Square Error) pour la minimisation locale par rapport à la fonction de densité réelle. L'objectif de cette partie est de montrer quelques techniques et laquelle utilisée pour la détermination du paramètre de lissage (basée sur les échantillons des données des durées de vie) pour que $\hat{f}_h(t) \to f(t)$.

Nous définissons l'erreur moyenne notée *MSE(t)* (Mean Square Error) en un point par rapport à la fonction densité réelle :

$$MSE(t) = E\left(\left(\hat{f}_h(t) - f(t)\right)^2\right), \tag{III.38}$$

en introduisant $E\left(\hat{f}_h(t)\right)$ dans l'expression du *MSE(t)*, nous obtenons

$$MSE(t) = E\left(\left(\hat{f}_h(t) - E\left(\hat{f}_h(t)\right) - f(t) + E\left(\hat{f}_h(t)\right)\right)^2\right), \tag{III.39}$$

le *MSE(t)* devient,

$$MSE(t) = E\left(\hat{f}_h(t) - E\hat{f}_h(t)\right)^2 + E\left(E\hat{f}_h(t) - f(t)\right)^2, \tag{III.40}$$

car l'espérance du produit croisé s'annule. Ainsi le *MSE* peut s'écrire en fonction du biais et de la variance de la fonction d'estimation :

$$MSE(t) = Var\left(\hat{f}_h(t)\right) + Bias\left(\hat{f}_h(t)\right)^2. \tag{III.41}$$

Une expression approximative du biais et la variance de l'estimateur de Parzen sont donnés par :

$$Biais(\hat{f}_h(t)) \approx \frac{1}{2}f"(t)h^2\int_{-\infty}^{\infty} u^2K(u)\mathrm{d}u + o(h^2), \tag{III.42}$$

avec

$$\sigma_K^2 = \int_{-\infty}^{\infty} u^2K(u)\mathrm{d}u, \tag{III.43}$$

$$Var(\hat{f}_h(t)) \approx \frac{1}{nh}f(t)\int_{-\infty}^{\infty} K^2(u)\mathrm{d}u + o\left(\frac{1}{nh}\right), \tag{III.44}$$

et

$$Rg(K) = \int_{-\infty}^{\infty} K^2(u)\mathrm{d}u. \tag{III.45}$$

Nous déduisons une approximation du MSE notée *AMSE* pour la sélection locale de h.

$$AMSE(t) = \frac{1}{4}f"(t)^2h^4\sigma_K^4 + \frac{1}{nh}f(t)Rg(K). \tag{III.46}$$

Le meilleur paramètre de lissage au sens du critère *AMSE* annule la dérivée de la fonction $AMSE$ pour un âge t donné. Cela implique,

$$h_{opt}(t) = \left(\frac{f(t)Rg(K)}{\sigma_K^4\left(f"(t)\right)^2}\right)^{1/5} n^{-1/5}. \tag{III.47}$$

Ce paramètre de lissage dépend de $f(t)$ et de sa dérivée seconde à l'instant t (inconnues en réalités). Il est inexploitable sans une expression paramétrique de la densité $f(t)$ et a une propriété locale. Par conséquent étant donné que nous cherchons à estimer une fonction sur

l'espace des temps, nous privilégions les paramètres de lissage résultant d'une optimisation globale. Les méthodes globales de calcul du paramètre de lissage. Les méthodes sont basées sur la minimisation d'une approximation de l'intégrale du carré l'erreur *ISE* ou sa moyenne *MISE*.

5.1 Méthodes basées sur *ISE*

Le paramètre de lissage est obtenu en minimisant l'intégrale de l'écart par rapport à la fonction que l'on veut estimer. Il est défini par :

$$\begin{aligned} ISE(h) &= \int_0^\infty \left(\hat{f}_h(t) - f(t)\right)^2 \mathrm{d}t, \\ &= \int_0^\infty \hat{f}_h(t)^2 \mathrm{d}t - 2\int_0^\infty \hat{f}_h(t) f(t) \mathrm{d}t + \int_0^\infty f(t)^2 \mathrm{d}t. \end{aligned} \tag{III.48}$$

Le terme $f(t)$ ne contenant pas le paramètre de lissage h n'apporte pas d'information sur celui-ci. Nous pouvons l'omettre de l'Équation (III.48). Ainsi, la minimisation l'*ISE* est identique à minimiser l'expression suivante notée *UCV* comme Unbiased Cross Validation.

$$UCV(h) = \int_0^\infty \hat{f}_h(t)^2 \mathrm{d}t - 2\int_0^\infty \hat{f}_h(t) f(t) \mathrm{d}t. \tag{III.49}$$

Le second terme à droite de l'égalité peut être estimé par la moyenne des $\hat{f}_h(t_i)$ d'après Stone [68]

$$E\int_0^\infty \hat{f}_h(t) f(t) \mathrm{d}t = \frac{1}{n}\sum_{i=1}^n \hat{f}_{-i}(t_i). \tag{III.50}$$

Où $\hat{f}_{-i}$ est obtenu par validation croisée comme ci-dessous.

$$\hat{f}_{-i}(t_i) = \frac{1}{(n-1)h} \sum_{j=1, i\neq j}^n K\left(\frac{t_i - t_j}{h}\right). \tag{III.51}$$

La technique *UCV* ou *LSCV*[1], la plus utilisée en pratique, a été étudiée par Bowman [69]. Bowman propose d'utiliser à la place du premier terme (voir Équation (III.48)) sa moyenne, par une estimation obtenue par validations croisées [69]. Nous obtenons en tenant compte de ce point, l'estimateur de densité avec validations croisées comme indiquée.

$$UCV(h) = \frac{1}{n}\sum_{i=1}^n \int_0^\infty \hat{f}_{-i}(t)^2 \mathrm{d}t - \frac{2}{n}\sum_{i=1}^n \hat{f}_{-i}(t_i). \tag{III.52}$$

1. Least Square Cross Validation

Le premier terme de l'expression à optimiser de Bowman est équivalent à

$$\int_0^\infty \hat{f}_h(t)^2 = \frac{1}{n}\sum_{i=1}^{n}\int_0^\infty \hat{f}_{-i}(t) + o\left(\frac{1}{n^2 h}\right). \tag{III.53}$$

à o près. Ce qui conduit à une expression proche de l'expression théorique de l'ISE.

$$LSCV(h) = \int_0^\infty \hat{f}_h(t)^2 \mathrm{d}t - \frac{2}{n}\sum_{i=1}^{n}\hat{f}_{-i}(t_i). \tag{III.54}$$

Cette nouvelle expression sera utilisées afin de trouver le paramètre de lissage au sens du critère *ISE*. La fenêtre de lissage est obtenue par minimisation, en utilisant la méthode des sécantes pour rechercher le paramètre h qui annule la dérivée première du *LSCV* ou en appliquant la méthode de bissection sur le *LSCV*. L'approximation de l'ISE par *LSCV* est plus plausible. Parce que son espérance mathématique à constante près est égale au *MISE*.

$$E(LSCV) = MISE + \int_0^\infty f(t)^2 \mathrm{d}t. \tag{III.55}$$

Le *LSCV* a des propriétés mathématiques intéressantes [70] car il constitue un estimateur sans biais à constante près (intégrale de la densité au carré) du *MISE*. Le paramètre de lissage issu converge en moyenne vers le paramètre optimale. Selon Hall et Marron [71] et [70] le rapport entre les paramètres de lissage $\left(\left(\frac{h_{lscv}}{h_{opt}}\right) - 1\right)^2$ a une distribuée Normale de moyenne nulle. Cependant le paramètre de lissage issue de *LSCV* a une grande variabilité [72] et [73] dépendant de l'échantillon ayant servi à déterminer le paramètre de lissage et possède au meilleur des cas un taux de convergence de l'ordre de $o\left(n^{-1/10}\right)$ [71]. En plus la recherche du paramètre de lissage optimal ne converge pas toujours [74]. D'après l'analyse de Sheater et Jones se basant sur le travail dans [75] l'estimation issue de la validation croisée a une forte variance due au fait que le paramètre de lissage issu est souvent plus petit que le paramètre optimal.

Devroye dans [56] a montré que la fonction d'estimation issue de cette méthode était non consistante pour certains types de fonction de densité $f(t)$. Drevoye montra que le paramètre de lissage issu de l'optimisation de mesure d'erreur absolue donne une borne inférieure en termes d'erreur absolue moyenne.

Feluch et Koronacki [72] ont amélioré la méthode du *LSCV* en supprimant en plus du terme $(i = j)$ les écarts les plus proche (c'est à dire $|t_i - t_j| < c_n$ avec $c_n > 0$). Le *TCV* (Timmed Cross Validation) proposée Feluch et Koronacki, est une amélioration de *LSCV* sous

2. h_{opt} paramètre de lissage optimal au sens du *MISE*

conditions que $f(t)$ soit quatre fois dérivable. Les auteurs [72] montrent que le paramètre de lissage issu du TCV est plus stable que celui de $LSCV$ et s'adapte plus aux distributions de densité multimodales et asymétriques.

$$TCV(h) = \int_0^\infty \hat{f}_h(t)^2 \mathrm{d}t - \frac{2}{n(n-1)h} \sum_{i=1}^{n} \sum_{j=1, j\neq i}^{n} K\left(\frac{t_i - t_j}{h}\right) 1_{|t_i - t_j| \geq c_n}. \tag{III.56}$$

Quant à nous pour l'usage de la méthode, nous proposons une intégration numérique pour obtenir le premier terme $\int_0^\infty \hat{f}_h(t)^2 \mathrm{d}t$ en discrétisant l'espace des données. Ces points seront pris différents des données de l'échantillon lors de l'optimisation de $LSCV$ au lieu d'utiliser les expressions approchées ou le TCV. Une intégration résultant du calcul de $f_h^2(t)$ en des points obtenus après une discrétisation de l'étendue de l'échantillon des durées de vie $\{t_i\}_{i=1,\cdots,n}$. Nous utiliserons la méthode d'intégration numérique de Simpson plus précise que la méthode du trapèze [21]. Ceci afin de réduire la variabilité de la procedure de recherche du paramètre de lissage.

5.2 Méthodes basées sur *AMISE*

AMISE est une approximation définie à partir du *MISE* ou l'intégrale de l'*AMSE*. Le critère se définit comme suite :

$$AMISE(h) = \frac{Rg(K)}{nh} + \frac{1}{4}\sigma_K^4 h^4 Rg(f"), \tag{III.57}$$

avec

$$Rg(g) = \int_{-\infty}^{\infty} g(u)^2 \mathrm{d}u, \tag{III.58}$$

et

$$\sigma_K^2 = \int_{-\infty}^{\infty} u^2 K(u) \mathrm{d}u. \tag{III.59}$$

La fenêtre optimale au sens du *AMISE* est donnée par le paramètre h qui annule sa dérivée relativement à h. Le paramètre de lissage optimal est :

$$h_{opt} = \left(\frac{Rg\,(K)}{\sigma_K^4 Rg\,(f")}\right)^{1/5} n^{-1/5}. \tag{III.60}$$

La valeur de $Rg(f")$ est inconnue et doit être estimée pour obtenir le paramètre de lissage. Les approches issues de l'*AMISE* recherchent une approximation $Rg(f")$. Un terme qui dépend de la fonction inconnue $f(t)$.

5.3 Méthode de validation croisée biaisée 1

Scott et Terrell ont proposé d'ajouter un biais à l'*AMISE* [73]. L'ajout du biais permet d'estimer la valeur de $Rg(f")$ et de réduire la variabilité du paramètre de lissage h. L'estimateur de la valeur de $Rg(f")$ est donné par :

$$\hat{R}g\,(f") = Rg\left(\hat{f"}_h\right) - \frac{Rg\,(K")}{nh^5}, \tag{III.61}$$

et

$$Rg\left(\hat{f"}_h\right) - \frac{Rg\,(K")}{nh^5} = \frac{2}{n^2h^5}\sum_{i=1}^{j}\sum_{j=1}^{n}\Theta_K\left(\frac{t_i - t_j}{h}\right). \tag{III.62}$$

où Θ_K est le produit de convolution défini comme suite :

$$\Theta_K(x) = \int_0^\infty K"(u)K"(u+x)du.$$

5.4 Méthode de validation croisée biaisée 2

Hall et Marrons [76] proposent un estimateur de $Rg(f")$ pour les noyaux très lisses se basant sur l'espérance de la dérivée quatrième.

$$\hat{R}g\,(f") = \frac{1}{n}\sum_{i=1}^{n} f_{-i}^{(4)}(t). \tag{III.63}$$

Une expression que les auteurs ont obtenu par intégration par partie du $Rg(f")$:

$$Rg\,(f") = \int_0^\infty f"(t)^2\mathrm{d}t, \tag{III.64}$$

appliquant deux intégrations par partie sur le terme de droite de l'égalité ci-dessus, nous obtenons :

$$\int_0^\infty f"(t)^2\mathrm{d}t = \int_0^\infty f^{(4)}(t)f(t)\mathrm{d}t, \tag{III.65}$$

d'où nous déduisons que $Rg(f")$ est équivalent à l'espérance mathématique de la dérivé d'ordre 4 de la fonction densité des durées de vie.

Ces méthodes d'estimation biaisées ont un ordre de convergence proche de celui de la méthode *LSCV(h)* [74]. Elles utilisent des dérivées d'ordre supérieures et ne sont applicables que lorsque nous avons des noyaux lisses (infiniment dérivable) tels que le noyau Normal. Nous remarquons qu'il est difficile de l'utiliser avec le noyau optimal d'Epanechnikov.

Les méthodes précédentes font aussi partie de la famille des méthodes de validation croisée

car elles n'utilisent pas le terme non-stochastique dans le calcul c'est à dire $(i = j)$. Elles ont été développées pour juste pallier la variabilité du paramètre de lissage de la méthode de $LSCV$ au tour de sa valeur moyenne.

5.5 Plug-in naïf

Les techniques basées sur le plug-in ont été introduites pour pallier la convergence lente des méthodes de validation croisée [70] et [74]. Elles consistent à proposer une estimation itérative et plus élaborée de $Rg(f")$ pilotée par une suite convergente de paramètres de lissage liée ou pas à h. Ces techniques sont itératives et inversent le problème de recherche de la valeur du paramètre de lissage. Elles accusent un degré de convergence plus rapide que les méthodes de validation croisée biaisée ou non biaisée [75] et avec une faible variance. Le plus difficile dans toutes les techniques basées sur l'$AMISE$ est la recherche d'estimateur robuste de $Rg(f")$.

Une technique du plug-in peut se réduire à la recherche du point fixe d'une fonction $g(X, h)$ en h où $X = \{t_1, t_2, \cdots, t_n\}$ le vecteurs de durée de vie et h le paramètre de lissage.

$$g(X, h) = \left(\frac{Rg\,(K)}{n\sigma_K^4 Rg\,(f")} \right)^{1/5}. \tag{III.66}$$

Le terme $Rg(f")$ est inaccessible, dû au fait que $f(t)$ nous est inconnue. De manière intuitive, nous pouvons utiliser une estimation naïve pour l'implémentation de la méthode du plug-in en prenant à chacune des étapes successives la dérivée seconde de l'estimateur de Parzen $Rg\left(\hat{f}_h"\right)$ pour estimer $Rg(f")$ comme indiqué dans l'Équation (III.67).

$$Rg(f") \approx Rg\left(\hat{f}_h"\right) = \frac{1}{n^2 h^6} \int_0^\infty \left(\sum_{i=1}^{n} K" \left(\frac{t - t_i}{h} \right) \right)^2. \tag{III.67}$$

Il suffit de résoudre l'Équation (III.68) en appliquant la méthode des sécantes afin d'obtenir la valeur du paramétre de lissage adéquate en fonction de l'erreur numérique tolérée.

$$h - \left(\frac{Rg(K)}{n\sigma_K^4 Rg\left(\hat{f}_h"\right)} \right)^{1/5} = 0. \tag{III.68}$$

Cette approche naïve implémentée ici pour le calcul des paramètres de lissage a des performances numériques supérieures à la classe des méthodes basées sur la validation croisée surtout pour les mélanges des données. Elle accuse quelques fois un problème de convergence causé par la dépendance directe de l'estimateur de $Rg(f")$ à h.

Dans [76], Hall et Marron exposent les deux classes d'estimateur $Rg(f")$. La première classe se base sur les produits de convolution et la seconde sur la dérivée d'ordre 4 de la fonction noyau en utilisant la validation croisée ou non.

5.6 Du Plug-in de Park & Marron à celui Sheater & Jones

Le plug-in de Sheater et Jones est une amélioration de la méthode proposée par Park et Marron dans [75] dont l'ordre de convergence est $n^{-4/13}$. Sheater et Jones [77] montrèrent qu'elle est la meilleure parmi l'ensemble des méthodes de calcul du paramètre de lissage avant de proposer une approche théoriquement plus performante, facilement programmable et plus précise.

L'estimateur de la fonction $Rg(f")$ pour le calcul du paramètre de lissage est similaire au cas de la méthode de validation croisée biaisée 2. Dans lequel nous estimons $Rg(f")$ en utilisant l'espérance mathématique de la dérivée d'ordre 4 de la fonction noyau. Contrairement au cas de la validation croisée biaisée l'estimateur de $Rg(f")$ n'est pas directement liée au paramètre de lissage cherché. Cet estimateur noté $S(\alpha(h))$, utilise une fonction noyau et des paramètres de lissage souvent distincts de celui utilisé dans $\hat{f}_h$. Dans leur travail Park et Marron considèrent l'estimateur sans biais de $Rg(f")$ qui est décrit par l'Équation (III.69) avec validation croisée (sans les termes non-stochastiques définis par $(i = j)$).

$$S(\alpha(h)) = \frac{1}{\alpha(h)^5 n(n-1)} \sum_{i=1}^{n} \sum_{j=1, i \neq j}^{n} K^{(4)} \left(\frac{t_i - t_j}{\alpha(h)} \right). \tag{III.69}$$

Avec $\alpha(h)$ comme le paramètre de lissage et $K(u)$ une fonction noyau lisse (Normal).
Sheater et Jones propose d'améliorer cette méthode (Park et Marron) par une injection de biais en supprimant la validation croisée dans $S(\alpha(h))$. Sheather et Jones ont montré dans [78] que l'ajout des termes non stochastique $(i = j)$ permettait d'accélérer la vitesse de convergence du paramètre calculé par plug-in vers le paramètre réel. Ceci avec une variation de variance négligeable par rapport au cas Park et Marron. Ils ont mis en place une démarche simple de calcul du paramètre de lissage qui conduisit à obtenir un ordre de convergence de $(n^{-5/14})$ vers

la solution optimale. Dans cette démarche $\alpha(h)$ est obtenu par calcul imbriqué :

$$S_{sj}\left(\alpha(h)\right) = \frac{1}{\alpha(h)^5 n(n-1)} \sum_{i=1}^{n}\sum_{j=1}^{n} K^{(4)}\left(\frac{t_i - t_j}{\alpha(h)}\right), \tag{III.70}$$

$$\alpha(h) = ch^{5/7},$$

$$c = 1.347\left(\frac{S_{sj}(a)}{T(b)}\right)^{1/7}, \tag{III.71}$$

$$T(b) = -\frac{1}{b^7 n(n-1)} \sum_{i=1}^{n}\sum_{j=1}^{n} K^{(6)}\left(\frac{t_i - t_j}{b}\right), \tag{III.72}$$

où a et b sont des paramètres de lissages pilotes définis à travers la distance interquartile des données IRQ :

$$a = 0.92 IRQ n^{-1/7}, \tag{III.73}$$

$$b = 0.912 IRQ n^{-1/9}, \tag{III.74}$$

Les méthodes de plug-in ont une grande performance par rapport aux autres techniques de calcul de paramètre [75], [77]. Les techniques du plug-in sont plus précises que la méthode de validation croisée non biaisée (vitesse de convergence de l'ordre de $o(n^{-1/10})$). Cependant, les méthodes plug-in en dépit de leur performance négligent la variation de l'estimateur $Rg(f")$. Car le paramètre de lissage issu de la dérivée de l'$AMISE$, est obtenu avec la fonction $Rg(f")$ qui est indépendante de h. Or son estimateur $S(\alpha(h))$ est lié au h. Sa variation en fonction du paramètre de lissage doit être prise en compte dans la minimisation de l'$AMISE$.

$$\frac{\partial}{\partial h} AMISE(h) = -\frac{Rg\,(K)}{nh^2} + \frac{1}{4}\sigma_K^4 \left(h^4 \frac{\partial}{\partial h} S((\alpha(h)) + 4h^3 S((\alpha(h))\right). \tag{III.75}$$

Le paramètre de lissage au sens de la méthode doit être le paramètre de lissage h qui minimise l'AMISE à travers une annulation l'Équation (III.75).

6 Applications numériques

6.1 Calcul du paramètre de lissage

Dans cette partie nous nous intéressons à l'efficacité statistique des méthodes de calcul du paramètre de lissage indispensable pour l'estimateur du Noyau de Parzen. Les données sont supposées provenir de distribution de Weibull. Une distribution de Weibull est définie par trois

paramètres dont :

- paramètre de forme β (scale parameter),
- paramètre d'échelle η (shape parameter),
- paramètre d'origine γ (location parameter).

Nous adopterons la notation suivante $W(\beta, \eta, \gamma)$ pour la distribution de Weibull dans le reste du rapport. La distribution de Weibull est définie par :

$$f(t) = \frac{\beta}{\eta}\left(\frac{t-\gamma}{\eta}\right)^{\beta-1} \exp\left(-\left(\frac{t-\gamma}{\eta}\right)^{\beta}\right), \tag{III.76}$$

$$F(t) = 1 - \exp\left(-\left(\frac{t-\gamma}{\eta}\right)^{\beta}\right), \tag{III.77}$$

$$R(t) = \exp\left(-\left(\frac{t-\gamma}{\eta}\right)^{\beta}\right), \tag{III.78}$$

$$\lambda(t) = \frac{\beta}{\eta}\left(\frac{t-\gamma}{\eta}\right)^{\beta-1}. \tag{III.79}$$

Chaque paramètre de la distribution de Weibull a une part d'explication dans le processus d'usure de l'équipement pendant son exploitation. Explicitement le paramère β décrit le mode de défaillance de l'équipement ou son intensité d'usure en fonction de l'âge. Le paramètre η désigne l'échelle ou la plage de variation de l'âge de fonctionnement de l'équipement. Le paramètre d'origine γ désigne l'âge au delà duquel l'équipement pourrait être sujet à des défaillances.

L'objective de ces simulations d'un côté est de comparer certaines techniques de calcul du paramètre de lissage sur les données de défaillance générées. D'un autre côté, elle permet de mettre en application la démarche d'estimation : des données de durées de vie jusqu'à la modélisation de la densité de probabilité. Les données sont supposées indépendantes et provenant de distributions de Weibull qui peuvent être de même mode ou non et même échelle ou non. Une hétérogénéité qui résulte du changement de mode, des conditions ou/et lieu de fonctionnement des équipements. Dans cette analyse, nous nous intéressons à la variabilité des calculs de paramètres de lissage issus des techniques citées dans les sections précedentes. Nous allons tester plusieurs types de distribution de Weibull telles que : distribution simple avec un seul Weibull, des mélanges de distributions de Weibull et des distributions de Weibull en concurrence.

Cas des distribution simples, les résultats sont représentés par des boites à moustache (boxplots) sur les Figures III.3, III.4 et III.5 respectivement pour les tailles d'échantillons 100, 500 et 1000. La boite à moustache est un outil graphique pour visualiser la dispersion dans un

échantillon autour de sa valeur médiane représentée par un trait rouge à l'intérieur de la boite. La longueur de la boite est définie par la distance entre le premier et le troisième quartiles. Les moustaches représentent les queues inférieure et supérieure. Les queues ou moustaches inférieure et supérieure sont définies respectivement par rapport au premier et troisième quartiles à 1,5 fois la longueur de la boite [3].

Nous notons qu'en fonction de l'augmentation de la taille les valeurs des paramètres de lissage et leurs dispersions respectives se réduisent. Nous remarquons en plus que les techniques n'ont pas les mêmes performances. Ainsi, la méthode du plug-in de Sheater et Jones (noté h_plg_ShJ_dich et h_plg_ShJ_grad) enregistre de faible dispersion contrairement aux paramètres calculés par la validation croisée baisée 2 (noté h_bcv2). A taille élevée, la méthode de validation croisée biaisée 1 (h_bc1) devient identique à la méthode de plugin naif (h_plgnaif) voir Figure III.5. Enfin, la méthode de validation croisée non biaisée (h_lscv) fournit des fenêtres de lissage souvent faibles comme l'indique les queues de son box-plot. En effet la méthode de validation croissée biaisée 2 doit être évitée vu sa dispersion et en plus elle ne converge pas souvent vers des solutions réelles.

Ces analyses sont identiques au cas de mélange pondéré de distributions de Weibull de même paramètre d'échelle voir Figures III.6, III.7 et III.8. Cependant, lorsque les paramètres d'échelle des distributions présentes dans le mélange sont significativement différents alors les autres techniques de calcul enrégistrent des insuffisances. Pour le mélange de trois distributions de Weibull de paramètre de d'échelle 15, 150 et 1500, nous remarquons que seul les méthodes du plug-in restent performantes et ne sont pas piégées par les données provenant de la distribution de Weibull de paramètre d'échelle 15. Par conséquent les Figures III.9, III.10 et III.11 montrent que les autres méthodes produisent des valeurs de paramètres de lissage faible. Cela conduit à une courbes d'estimation rugueuse, signe d'une forte variance d'estimation. Nous notons également que les validations croisées non biaisée et biaisée 2 ont des performances très faibles en mélange des données lorsque les paramètres d'échelle ne sont pas du même ordre de grandeur.

Pour le cas de la concurrence, l'ordre des performances demeurent identiques en terme de variabilité des résultats. Cependant, il faut remarquer que la méthode du Noyau de Parzen est une approche fréquentielle alors que la densité issue de la concurrence ne l'est pas. Lorsque les données ne sont pas linéairement séparable c'est à dire $\beta_1 << \beta_2$ ou $\eta_1 << \eta_2$ la méthode du noyau n'est pas apte à faire un lissage correcte de la densité de probabilité des durées de vie et toutes les autres fonctions qui en découlent.

3. dans les graphiques la moustache inférieure est définie par la plus petite des observations supérieure au premier quartile diminué de 1,5 fois la longueur de la boite et la moustache supérieure est définie par la plus grande des observations inférieure au troisième quartile augmenté de 1,5 fois la longueur de la boite

Enfin, les méthodes basées sur les fonctions de répartition ont un temps de calcul énorme dû au calcul d'intégrations numériques. Les paramètres résultants de la minimisation du $CE(h)$ (voir annexe) sont trop large conduisant à un biais très élévé Figure III.15 ou III.16. Les autres basées sur le critère $EC(h)$ (voir annexe) sous estiment le paramètre de lissage avec un temps de calcul qui explose avec la taille des données. Cependant, cette sous-estimation dépend du facteur de p utilisé dans le calcul de la distribution empirique de la fonction de répartition (ici nous avons utilisé $p = \{0; 0,1; 0,5; 0,7; 0,9\}$ correspondant respectivement aux paramètres de lissage h_rep0, h_rep1, h_rep5, h_rep7, h_rep9).

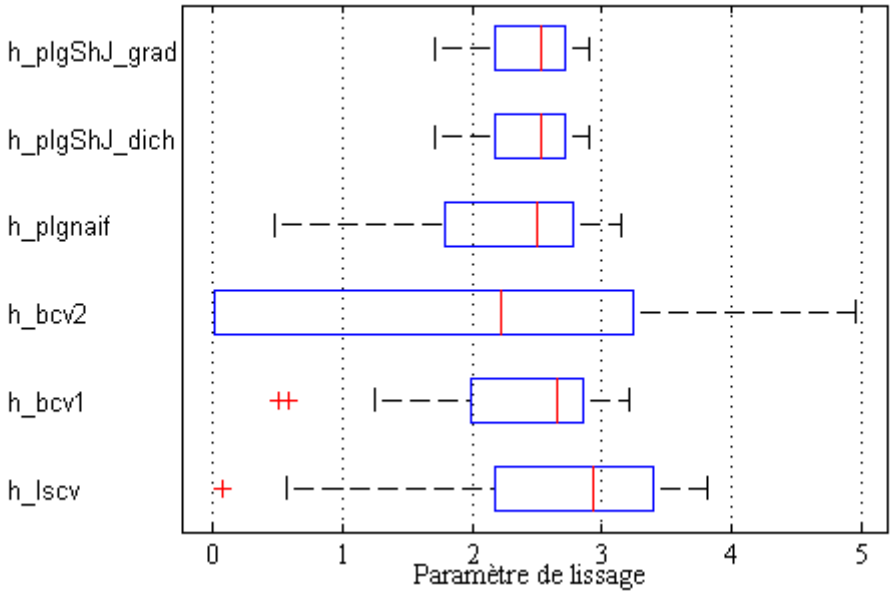

Figure III.3 – Dispersion après 30 simulations de calcul du paramètre de lissage h : cas d'une Weibull $W(2,3;15)$; taille des données 100.

6.2 Analyse des résultats

A travers les simulations et sur la base des résultats des études théoriques, quelques auteurs encouragent le choix les méthodes basées sur le plug-in pour le calcul du paramètre lissage et cela quel que soit le type des données. Ce choix se justifie par la capacité d'apprentissage dans le cas des données complexes de la technique du plug-in. À travers les figures, nous notons une faible dispersion de ses paramètres de lissage. Les paramètres de lissage issus de la méthode de validation croisée biaisée 2 sont très dispersés. Les résultats statistiques de la méthode de validation biaisés 1 sont proche de ceux du plug-in simple/naïf. La méthode de validation croisée

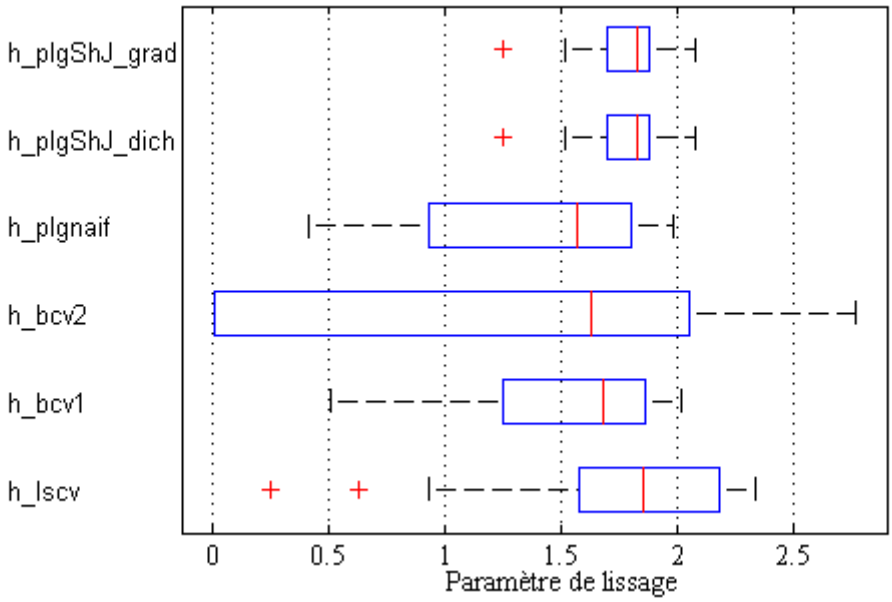

Figure III.4 – Dispersion après 30 simulations de calcul du paramètre de lissage h : cas d'une Weibull $W(2,3;15)$; taille des données 500.

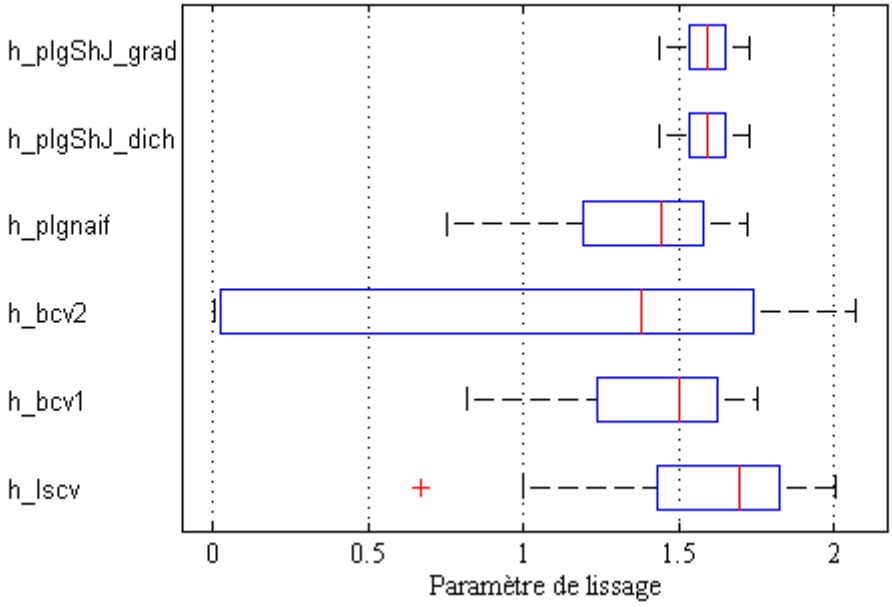

Figure III.5 – Dispersion après 30 simulations de calcul du paramètre de lissage h : cas d'une Weibull $W(2,3;15)$; taille des données 1000.

biaisée 2 est la moins efficace pour ce type de distribution surtout à grande taille. Les méthodes de plug-in Sheater et Jones avec gradient enregistrent une faible dispersion. Par conséquent sur

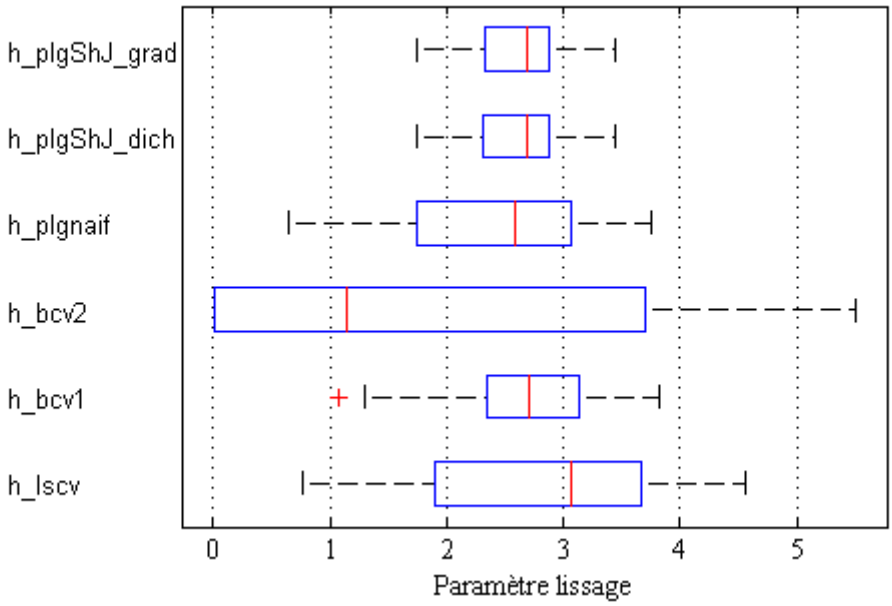

Figure III.6 – Dispersion après 30 simulations de calcul du paramètre de lissage h : cas de mélange pondéré de Weibull de même paramètre d'échelle $(35\%W(1,5;15) + 55\%W(2,3;15) + 10\%W(4,5;15))$; taille des données 100.

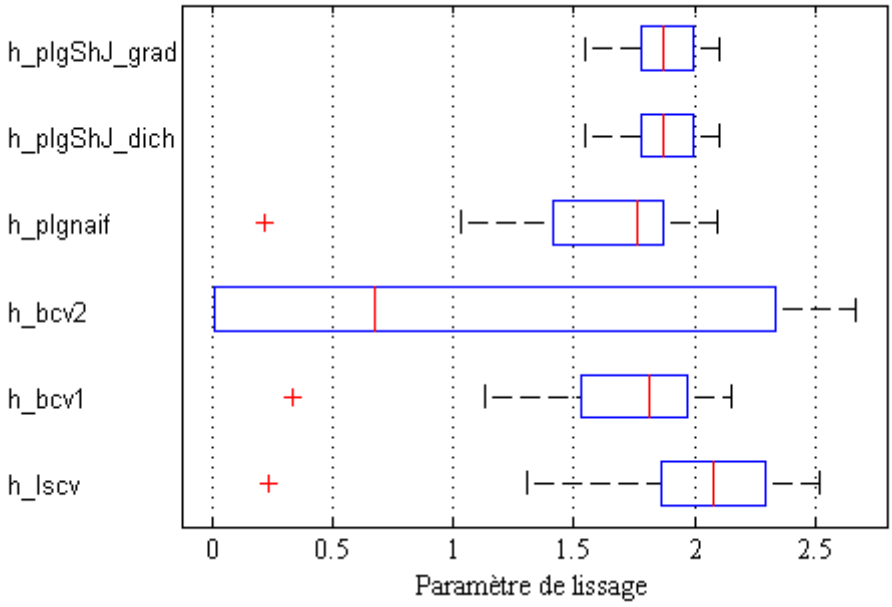

Figure III.7 – Dispersion après 30 simulations de calcul du paramètre de lissage h : cas de mélange pondéré de Weibull de même paramètre d'échelle $(35\%W(1,5;15) + 55\%W(2,3;15) + 10\%W(4,5;15))$; taille des données 500.

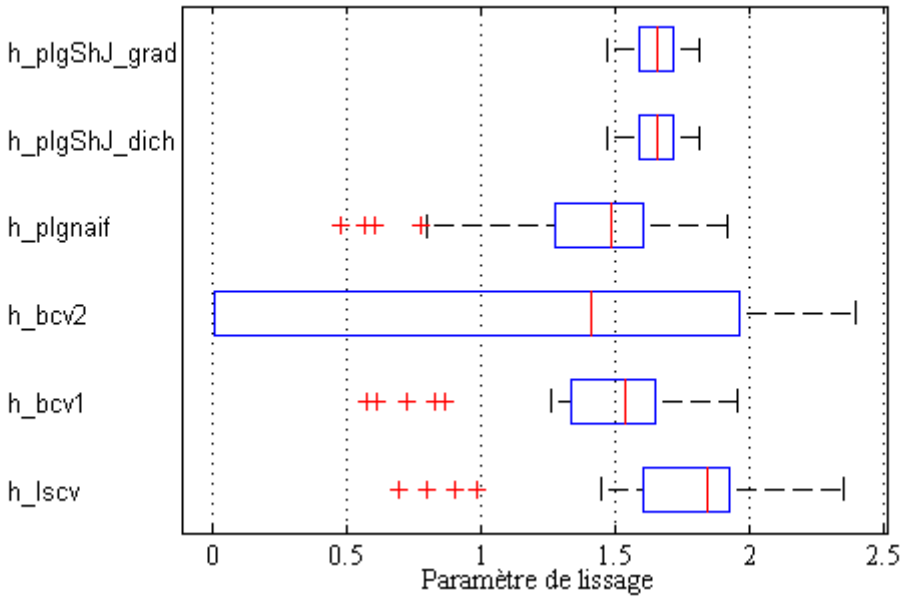

Figure III.8 – Dispersion après 30 simulations de calcul du paramètre de lissage h : cas de mélange pondéré de Weibull de même paramètre d'échelle $(35\%W(1,5;15) + 55\%W(2,3;15) + 10\%W(4,5;15))$; taille des données 1000.

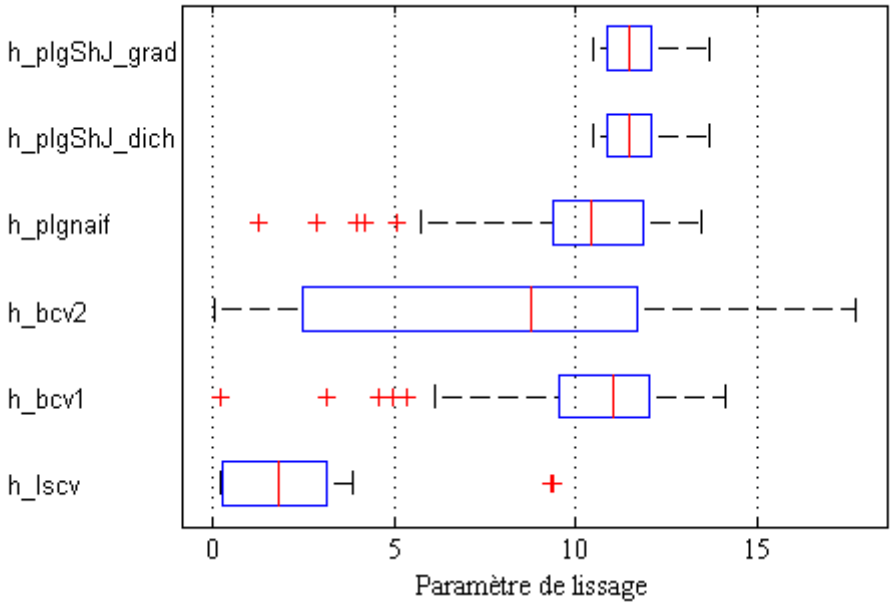

Figure III.9 – Dispersion après 30 simulations de calcul du paramètre de lissage h : cas de mélange pondéré de Weibull de même paramètre de forme $(35\%W(2,5;15) + 55\%W(2,5;150) + 10\%W(2,5;1500))$; taille des données 100.

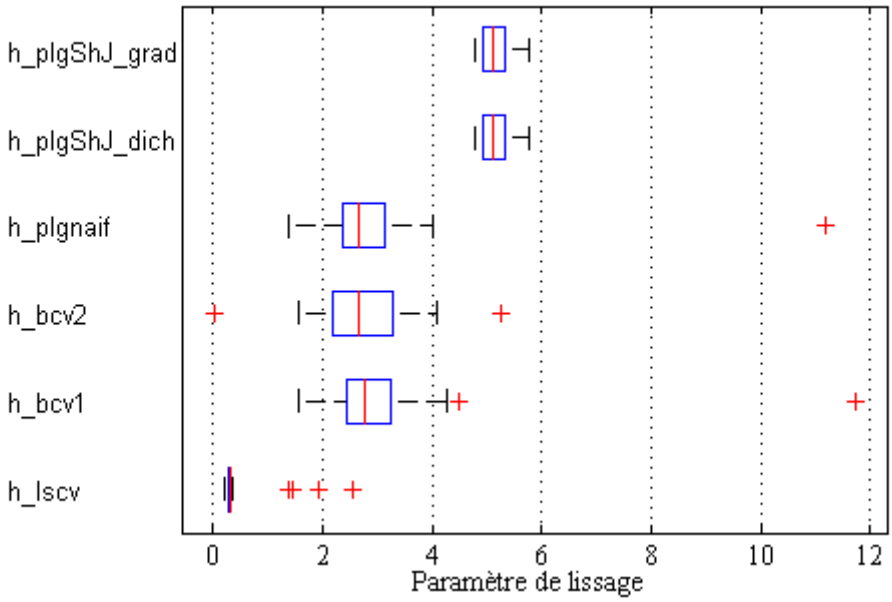

Figure III.10 – Dispersion après 30 simulations de calcul du paramètre de lissage h : cas de mélange pondéré de Weibull de même paramètre de forme $(35\%W(2,5;15)+55\%W(2,5;150)+10\%W(2,5;1500))$; taille des données 500.

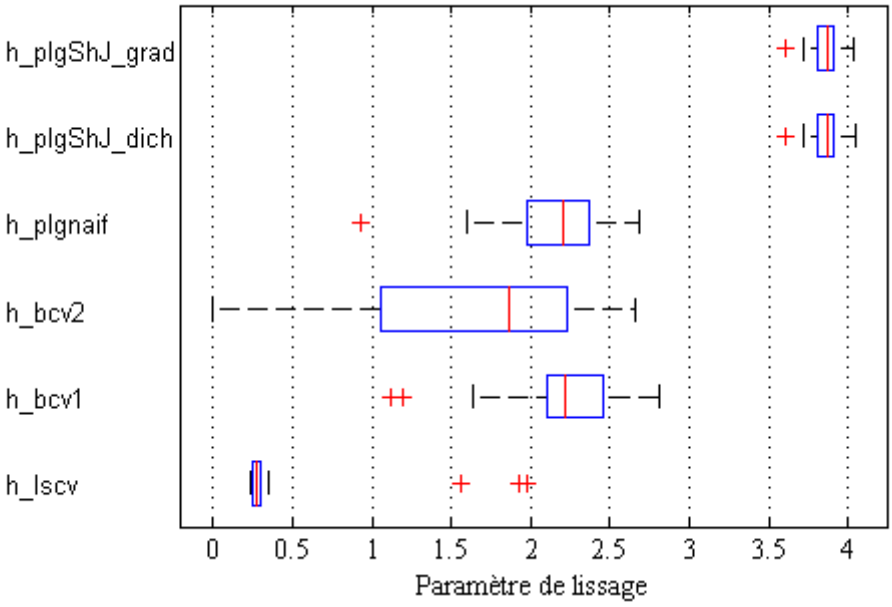

Figure III.11 – Dispersion après 30 simulations de calcul du paramètre de lissage h : cas de mélange pondéré de Weibull de même paramètre de forme $(35\%W(2,5;15)+55\%W(2,5;150)+10\%W(2,5;1500))$; taille des données 1000.

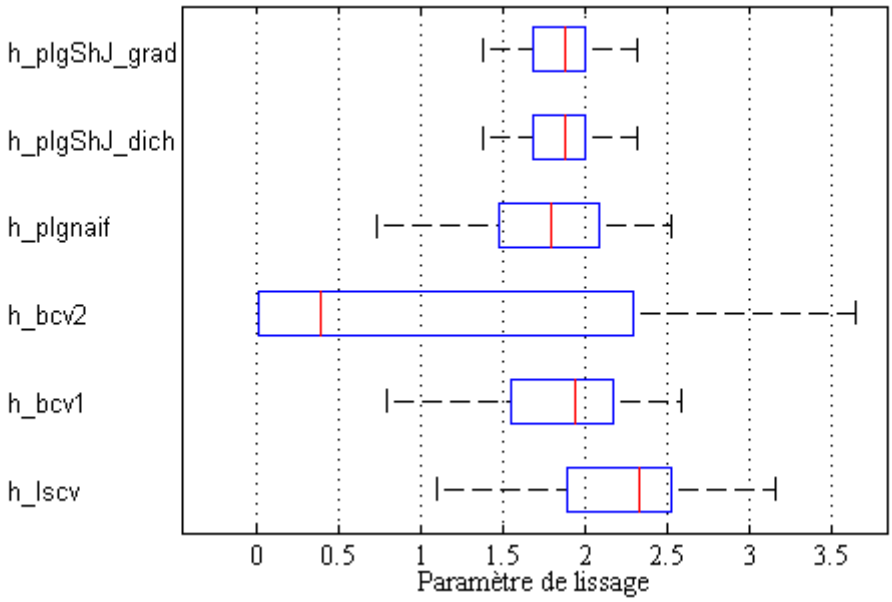

Figure III.12 – Dispersion après 30 simulations de calcul du paramètre de lissage h : cas des données provenant d'un équipement composé de sous-équipements en concurrence $(Min\{W(1,5;15);W(2,3;15)\})$; taille des données 100.

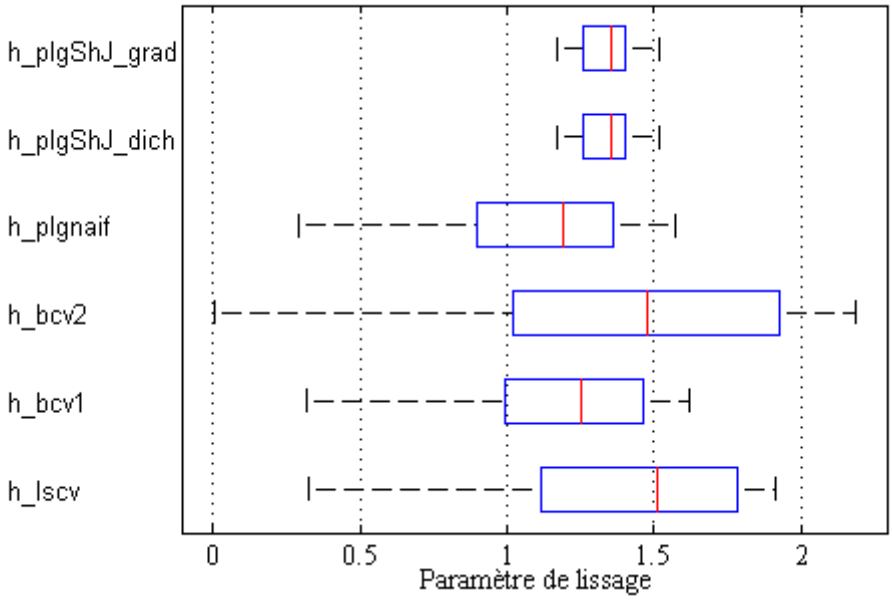

Figure III.13 – Dispersion après 30 simulations de calcul du paramètre de lissage h : cas des données provenant d'un équipement composé de sous-équipements en concurrence $(Min\{W(1,5;15);W(2,3;15)\})$; taille des données 500.

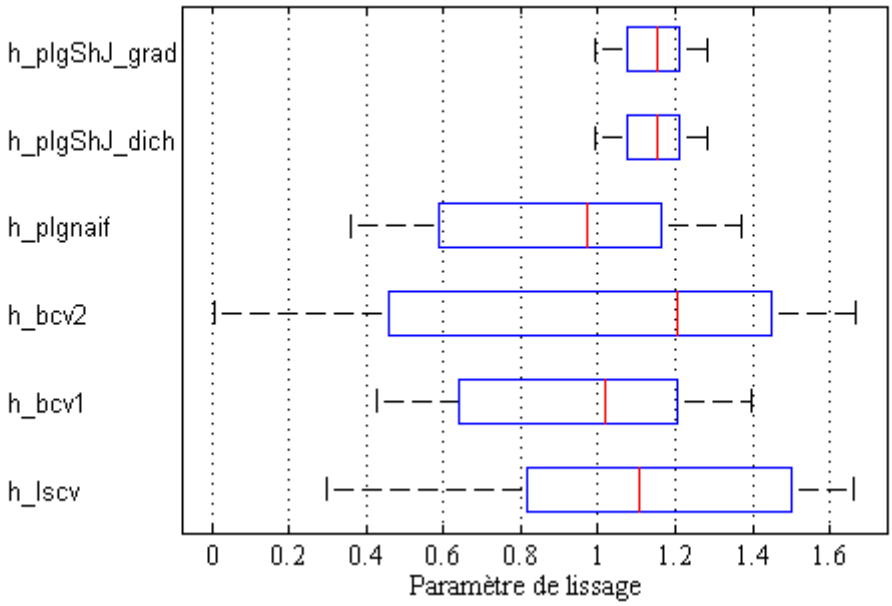

Figure III.14 – Dispersion après 30 simulations de calcul du paramètre de lissage h : cas des données provenant d'un équipement composé de sous-équipements en concurrence $(Min\{W(1,5;15);W(2,3;15)\})$; taille des données 1000.

ce critère, la méthode de plug-in de Sheater et Jones a de meilleures aptitudes pour l'estimation des densités.

Nous notons qu'il existe d'autre approche de calcul du paramètre de lissage telles que la méthode graphique de Silverman[44] ou la Smoothed Cross validation Method dans la littérature. La comparaison entre les méthodes est assez difficile et nous met face à un problème de généralisation. C'est à dire que nous ne pouvons pas prendre les résultats d'une simulation particulière et les généraliser. Cependant, nous notons théoriquement que la méthode du plug-in à un taux de convergence de l'ordre de $n^{-4/15}$ le meilleur parmi toutes les méthodes et le plus proche de la valeur théorique possible $n^{-1/2}$.

L'estimateur du noyau de Parzen, en dépit de son aptitude à modéliser des distributions complexes, montre quelques problèmes lorsque le support des données est borné selon Hominal et Deheuvels [42] pour l'estimation des fonctions de densité. C'est pour cette raison que nous nous sommes abstenus de prendre des données issues de distribution de Weibull à faible paramètre de forme β dans lesquels l'effet du biais est important au voisinage de 0. Ceci relate un problème de cette approche qui ne se résume pas seulement au seul choix du paramètre de lissage. L'application numérique suivante discute sur les solutions proposées dans la littérature afin d'améliorer la mise place de la démarche du noyau de Parzen et son utilisation.

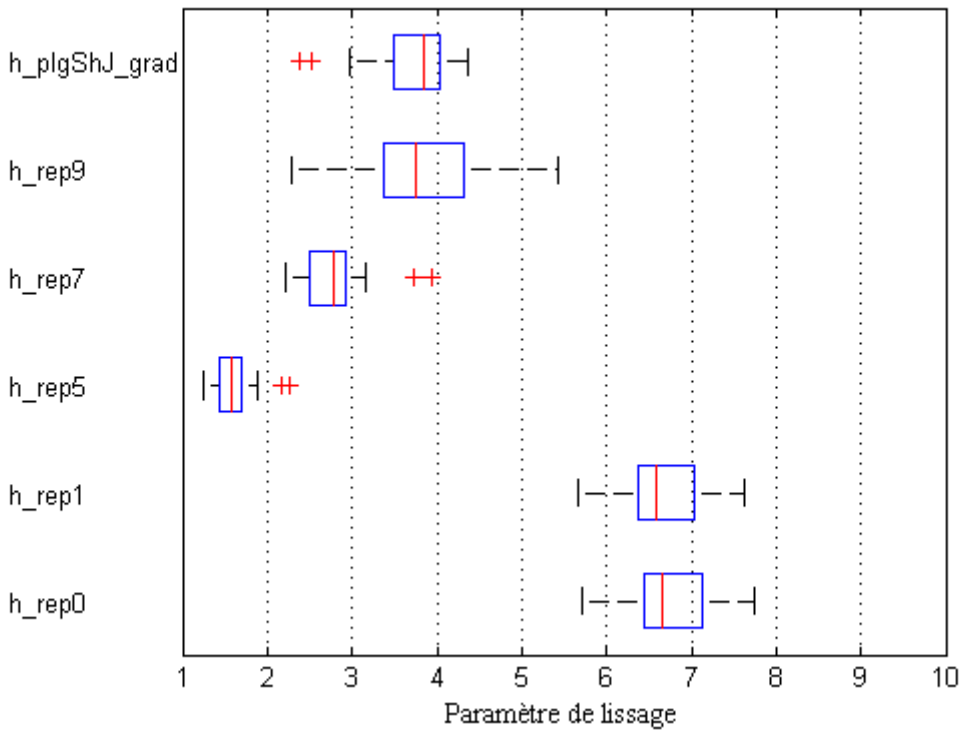

Figure III.15 – Dispersion après 30 simulations de calcul du paramètre de lissage h : cas basé sur les fonctions de répartition et des données provenant d'une distribution de melange de Weibull $(35\%W(2,5;15) + 55\%W(2,5;150) + 10\%W(2,5;1500))$; taille des données 100.

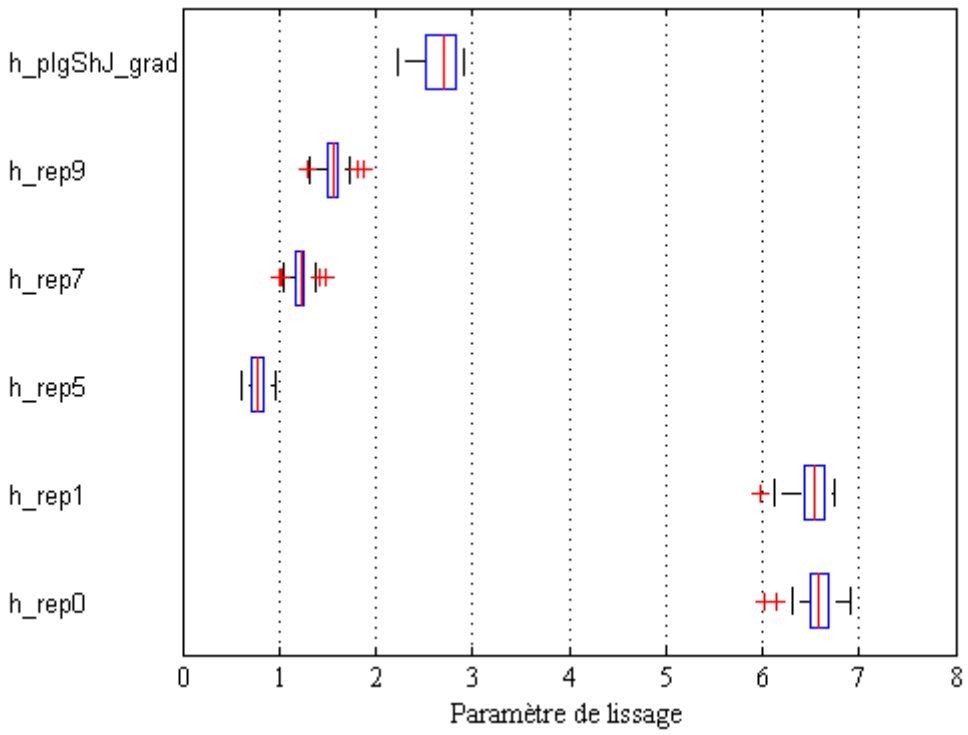

Figure III.16 – Dispersion après 30 simulations de calcul du paramètre de lissage h : cas basé sur les fonctions de répartition et des données provenant d'une distribution de melange de Weibull $(35\%W(2,5;15) + 55\%W(2,5;150) + 10\%W(2,5;1500))$; taille des données 500.

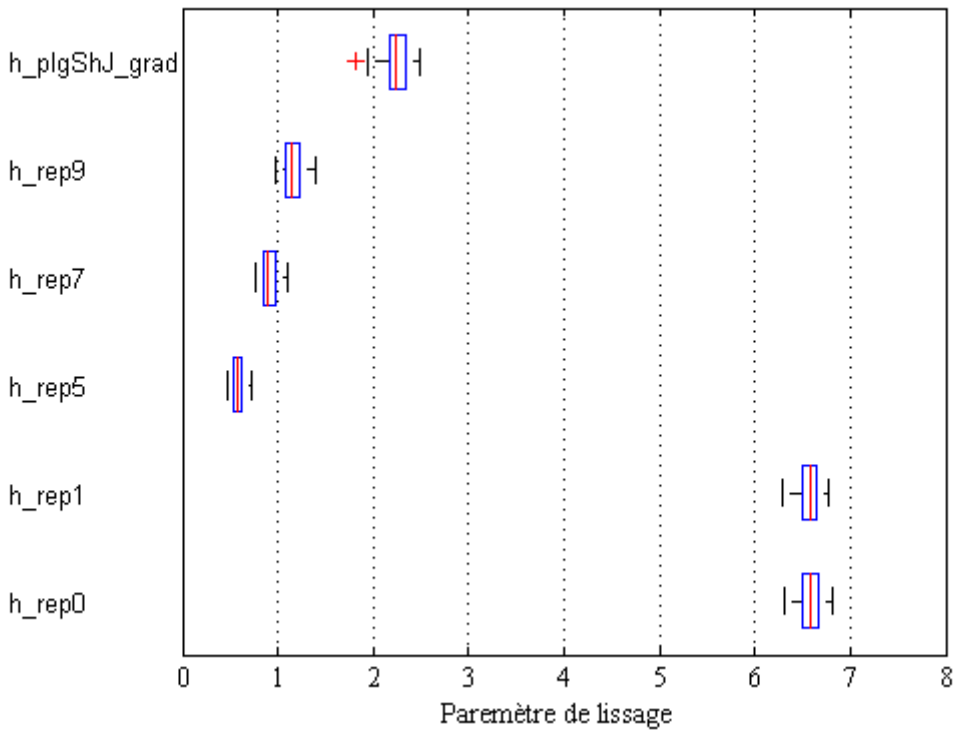

Figure III.17 – Dispersion après 30 simulations de calcul du paramètre de lissage h : cas basé sur les fonctions de répartition et des données provenant d'une distribution de melange de Weibull $(35\%W(2,5;15) + 55\%W(2,5;150) + 10\%W(2,5;1500))$; taille des données 1000.

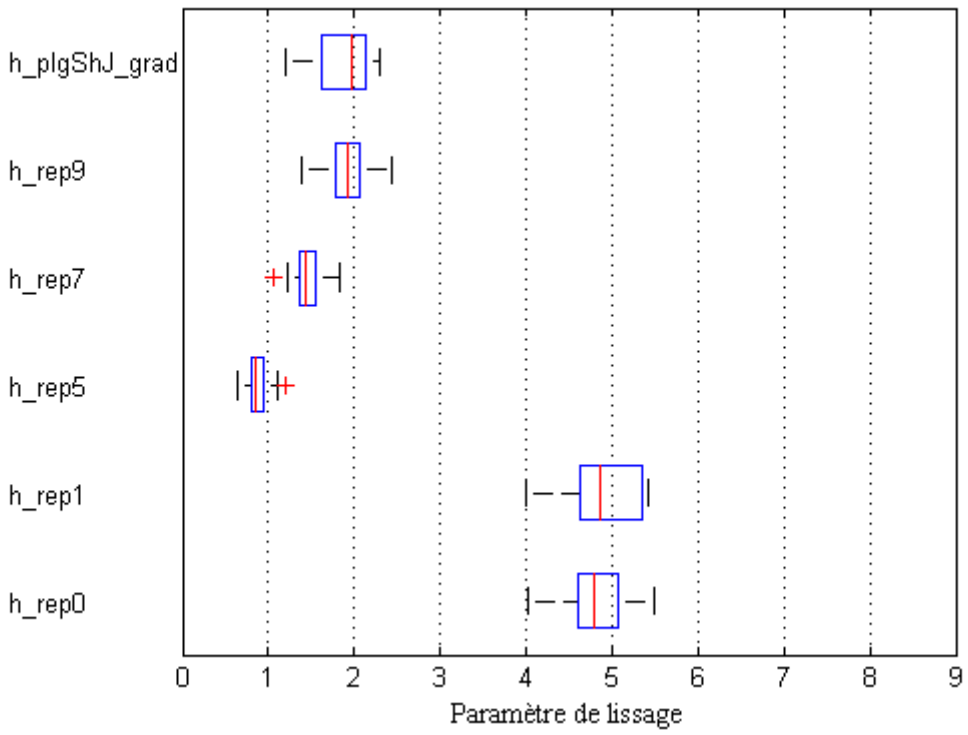

Figure III.18 – Dispersion après 30 simulations de calcul du paramètre de lissage h : cas basé sur les fonctions de répartition et des données provenant d'un équipement composé de sous-équipements en concurrence $(Min\{W(1,5;15);W(2,3;15)\})$; taille des données 100.

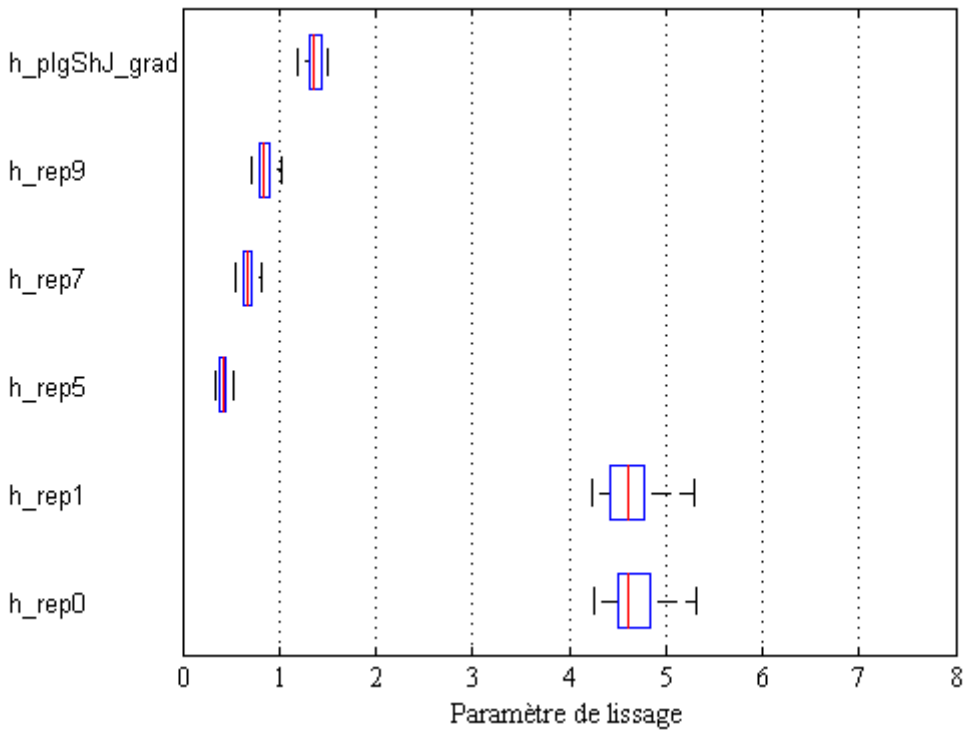

Figure III.19 – Dispersion après 30 simulations de calcul du paramètre de lissage h : cas basé sur les fonctions de répartition et des données provenant d'un équipement composé de sous-équipements en concurrence $(Min\{W(1,5;15);W(2,3;15)\})$; taille des données 500.

6.3 Correction de l'effet de bord

Dans cette partie, nous cherchons à modéliser la densité de panne des données provenant d'une distribution de Weibull $W(0,8;0,1)$ d'une taille $N = 500$. Pour cela, nous testons les méthodes de Marron et Ruppert [46] Figure III.20 et la transformation de Saoudi et al.,[43] Figure III.21 . Ces techniques de transformation ont un aspect global et proposent de meilleurs résultats en dépit de la complexité dans leur déroulement. Les autres, ayant un trait local, sont peu efficaces surtout face au mélange des données de durée de vie. Cependant, l'inefficacité de la méthode du Noyau de Parzen n'est importante que pour les distributions de Weibull pour lesquelles le paramètre de forme est faible c'est à dire $\beta \in]0;1,1[$ (où au voisinage de zéro la valeur de la dérivée $|f'(0)|$ peut être importante). Autrement, pour les distributions de Weibull dont le paramètre de forme β est supérieur à 2, nous pouvons se passer de ces différentes corrections et utiliser uniquement un des algorithmes de plug-in.

Parmi la classe des transformations des données, nous utiliserons la technique développée par Saoudi et al.,[43] pour les cas de fortes présences de biais au détriment de celle proposée par Marron et Ruppert[46]. La seconde technique est plus complexe et ne s'adapte pas directement au mélange des données.

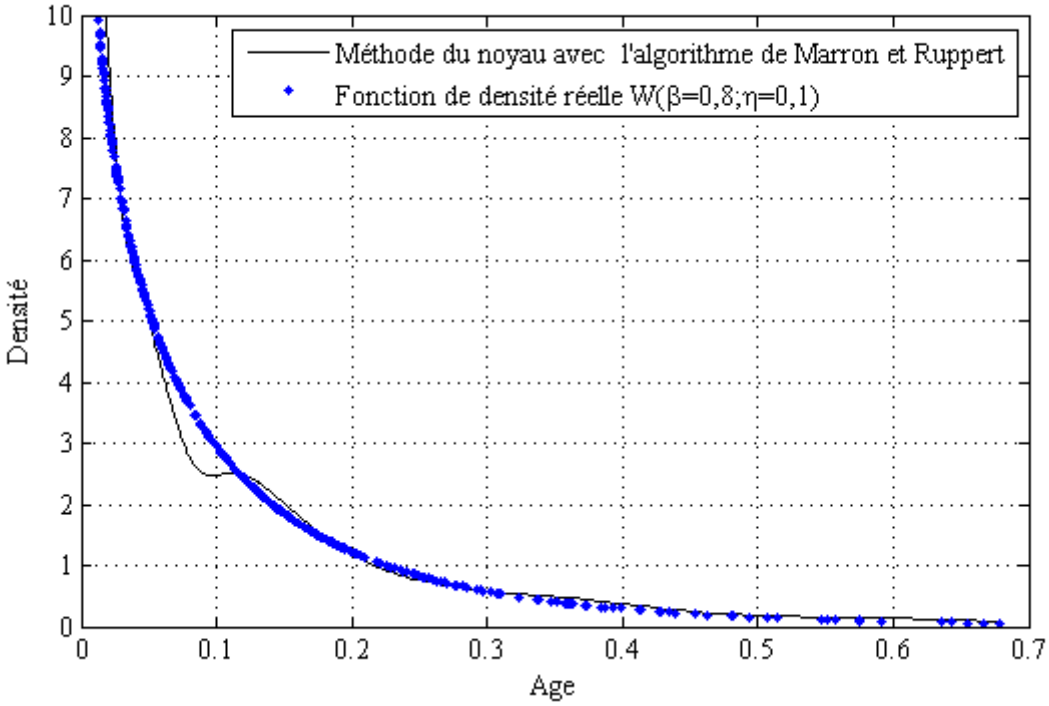

Figure III.20 – Estimation par la transformation de Marron et Ruppert cas $\{W(0,8;0,1)\}$ taille des données 500.

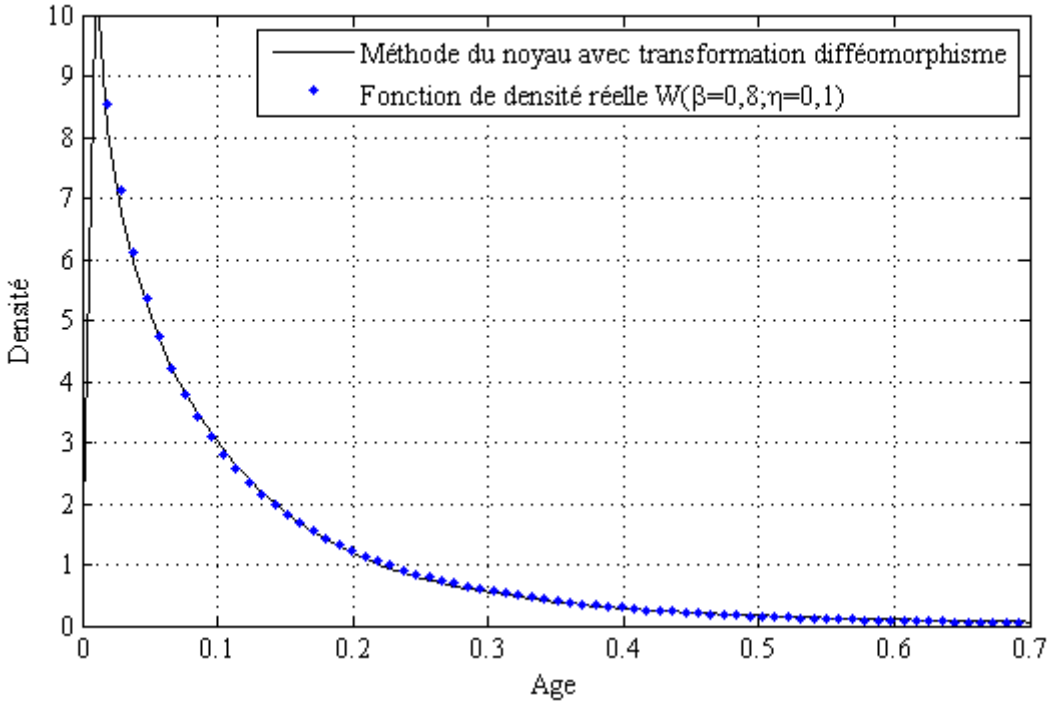

Figure III.21 – Estimation par la transformation difféomorphisme cas $\{W(0,8;0,1)\}$ taille des données 500.

7 Conclusion

La méthode du Noyau de Parzen a fait l'objet d'étude de ce chapitre. Nous avons exposé ses propriétés mathématiques et son aptitude à modéliser les données de durées de vie hétérogènes tout en réduisant la propagation des erreurs provenant du choix et de la variabilité de ses paramètres de lissage et de la fonction noyau.

Dans ce chapitre, plusieurs techniques de calcul des paramètres de la méthode du Noyau de Parzen furent décrites et implémentées. Nous avons également programmé plusieurs techniques de calcul du paramètre de lissage. Par la suite, nous avons réalisé une analyse comparative entre les différentes techniques sur le critère de variabilité des résultats de calcul des paramètres de lissage. En prenant la dispersion comme critère de sélection, nous avons choisi la technique du plug-in pour le calcul du paramètre de lissage. Cette technique de calcul présente le taux de convergence le plus rapide de toutes les techniques présentées dans ce rapport. De plus, nous avons montré que les estimations basées sur le paramètre de lissage issu de la technique du plug-in, s'adaptent plus facilement aux données de durées de vie hétérogènes.

Nous avons terminé ce chapitre par présenter les insuffisances de la méthode du Noyau de Parzen lorsque les supports des données sont semi-bornés ou bornés. Ces insuffisances se traduisent par de forts biais au voisinage des bords. Ce chapitre présenta également une revue sur les approches utilisées pour réduire un tel effet. Ces approches se regroupent en approches locales et globales. Les approches locales sont simples à mettre en œuvre et se basent sur la réflexion des données de durées de vie ou encore sur la renormalisation. Cependant, ces approches restent peu efficaces. Par contre, les approches ayant un caractère global, utilisent les noyaux asymétriques ou la transformation des données de durées de vie. Ces dernières approches sont plus efficaces que les approches locales en dépit de la complexité de leur mise en œuvre. Nous avons opté pour l'usage de la transformation logarithmique des données de durées de vie pour réduire le biais de bord dans nos estimations.

La méthode d'estimation du noyau expliquée, ici, sera exploitée dans le chapitre suivant pour proposer un modèle statistique des coûts engendrés par des politiques de maintenance. L'analyse dans le prochain chapitre portera essentiellement sur l'étude des politiques de maintenance appliquées à des équipements usagés et évoluant dans plusieurs milieux distincts. Dans cette analyse les distributions des durées de vie des équipements seront supposées inconnues et estimées par la méthode du Noyau de Parzen.

Chapitre IV

Analyse de la polique de maintenance de type âge pour un équipement usagé

Ce chapitre propose un modèle statistique et l'analyse de la politique de maintenance de type âge appliquée à un équipement ayant été exploité dans un environnement de caractéristique bien spécifique. Dans cette analyse, contrairement à la majorité des approches existantes, la fonction de distribution des durées de vie est supposée inconnue. Dans ce présent travail, l'équipement est caractérisé seulement par des données des durées de vie. La fonction de distribution ainsi que les fonctions objectif sont estimées par la méthode du Noyau de Parzen.

L'objectif de ce chapitre est double. La première partie de ce chapitre analyse théoriquemment la politique de maintenance de type âge pour les équipements évoluant concécutivement dans des environnements d'exploitation distincts. Dans l'analyse théorique, nous prouvons l'existence et l'unicité d'un âge optimale de remplacement. Deuxièmement, une étude statistique est réalisée dans laquelle nous prouvons la faible variabilité des résultats issus des estimations avec la taille des données. Des extensions du modèle proposé seront aussi théoriquement étudiées.

1 Introduction

L'intérêt croissant pour une meilleure gestion des équipements industriels a toujours suscité de pertinents travaux sur la modélisation et l'analyse mathématiques des coûts d'exploitation. Un équipement, pour produire de la valeur, se dégrade en fonction de l'environnement et de la sollicitation. Sans une politique de maintenance préventive adéquate, l'équipement serait fréquemment sujet à des pannes. Ces pannes d'équipement engendrent des coûts non-planifiés et des risques dont les conséquences peuvent être grave pour l'environnement et les agents d'exploitation.

Plusieurs travaux intéressants ont été réalisés dans la littérature sur l'analyse et la mise en place de politiques de maintenance. Ces politiques de maintenance ont largement été discutées dans [1] [2] [79] par Barlow, Hunter, Proschan, Rosenblatt J.R, Weiss et Wolman, pour les équipements non-réparables. Depuis, plusieurs extensions ont été réalisées incluant les équipements réparables [34] [80] [81]. Dans [82], Nakagawa et Mizutani ont étudié le remplacement périodique avec une réparation minimale sur un horizon fini. Cette approche est un développement de la politique de maintenance de type block classique étudiée sur un horizon d'optimisation infini dans [2]. Pour plus de détail sur les politiques de maintenance préventives sur horizon fini, nous pouvons se référer aux travaux de Nakagawa [83].

Tous les travaux cités ci-dessus reposent sur l'hypothèse que la distribution de durées de vie des équipements appartient à une des familles paramétriques : Gamma, Weibull ou Normale tronquée $\cdots$. Les modèles paramétriques sont statistiquement validés sur la base des données de durées de vie. Les travaux de Lawless [13] montrent que plusieurs modèles paramétriques peuvent être statistiquement validés pour un échantillon de données. Dans [13], Lawless montre sur un exemple réel que l'inférence avec deux modèles testés et statistiquement validés peut être complètement différente. Par conséquent à l'erreur due aux choix, parmi les modèles paramètriques, nous ajoutons les erreurs rélatives à l'estimation de ses paramètres sur la base d'un échantillon. Alors, l'approche paramétrique introduit deux sources d'erreur dans l'optimisation de la maintenance basée sur les données de durées de vie d'où l'utilisation des méthodes d'estimation non-paramétrique pour réduire le nombre de sources d'erreurs. Ce present chapitre étudie la politique de maintenance de type âge pour un équipement subissant des pannes. L'équipement est supposé avoir été exploité dans un premier environnement. À la fin de la mission dans cet environnement, l'âge de l'équipement est egal à T_1. Le même équipement est conduit dans un autre environnement différent du premier pour une nouvelle mission. Les conditions d'exploitation dans le nouvel environnement peuvent être plus ou moins sevère que dans le précedent. Cette différence est modélisée par des lois de probabilité distinctes rélative-

ment aux conditions d'exploitation. Afin de réduire les défaillances pendant l'exploitation dans le second environnement l'équipement subit une politique de remplacement de type âge. Sous cette politique de maintenance l'équipement est remplacé à la panne (maintenance corrective) ou après un certain âge T sans défaillance (maintenance préventive).

Le travail consiste ici à déterminer l'âge optimal T_0 pour réaliser la maintenance préventive garantissant un coût de maintenance minimal. Ce coût de maintenance tient compte du coût de la maintenance préventive et corrective. Le modèle du coût est estimé par la méthode du Noyau de Parzen sur la base des données de durées de vie recoltées dans les mêmes conditions d'exploitation.

Nous faisons la remarque qu'il existe des travaux antérieurs sur l'analyse non-paramétrique de la politique de maintenance de type âge pour les équipements neufs. Parmi lesquels nous pouvons citer ceux de Coolen et ses co-auteurs [84] [85]. Pour l'estimation du coût maintenance, Coolen et ses co-auteurs utilisent une estimation empirique de la fonction de distribution de durées de vie de l'équipement. Le modèle de Coolen et ses co-auteurs présente des discontinuités dues à l'utilisation d'estimateur empirique de la distribution des durées de vie. Cette discontinuité pousse les auteurs à donner un encadrement du coût de la maintenance. Les âges optimaux pour la réalisation de la maintenance préventive sont alors déduits de la minimisation des bornes supérieure et inférieure du coût maintenance.

Bergman et Klefsjo [86] [87] proposent une méthode graphique sur le concept du Temps Total de Test TTT pour resoudre le même problème de politique de maintenance de type âge pour les équipements neufs. Pour plus de détail sur le concept TTT nous pouvons se référer aux travaux de Barlow et Campo [88] ou Bergman [87]. De manière similaire, Dohi et ses co-auteurs [89] [90] [91] ont resolu le problème de temps limite de réparation en utilisant le même concept. Dans ces travaux, Dohi et ses co-auteurs [89] [90] [91] proposent une méthode pour déterminer le temps limite acceptable pour la réparation par rapport au temps de remplacement par du neuf. D'un autre côté Rinsaka et Dohi [92] combine le concept TTT avec la méthode du Noyau de Parzen pour optimiser la disponibilité des logiciels informatiques. Dans ce travail, Rinsaka et Dohi [92] ont prouvé que la méthode du Noyau de Parzen est efficace pour modéliser le coût de maintenance.

Notre approche est complètement basée sur les données de durées de vie et propose une estimation continue du coût de maintenance pour la politique de maintenance de type âge, contrairement à l'approche de Coolen et Coolen [85]. Une telle modélisation évite une estimation par intervalle de la fonction coût comme proposé dans [85]. De plus avec la méthode d'estimation du Noyau de Parzen, l'analyse de sensibilité de la fonction coût de maintenance devient plus facile. Dans notre travail la robustesse du modèle est montrée relativement aux choix et variation

des paramètres de l'estimateurs du Noyau de Parzen. Le chapitre IV s'organise autour de trois sections. La Section 2 pose le problème et donne une formulation probabiliste du coût unitaire de la politique de maintenance adoptée. Nous proposons, dans la même section, une estimation du coût probabiliste unitaire par la méthode du Noyau de Parzen. Un exemple numérique est fourni dans la Section 3 où nous comparons les résultats issus des modèles probabilistes et statistiques. Dans cette même section, nous analysons l'évolution du coût et l'âge optimaux de remplacement en fonction de la durée de la mission dans le premier environnement. Nous terminons le chapitre par la Section 4. Dans cette section, nous traitons les extensions du modéle de base défini dans la Section 2. Dans ces extensions, nous prenons en compte les coûts de maintenances antérieures exécutées dans le premier environnement d'exploitation.

2 Modèle mathématique de la politique de maintenance

Nous considérons un équipement sujet à des défaillances aléatoires. Pour diminuer le risque d'occurrence des défaillances une politique de maintenance préventive est adoptée : selon cette politique de maintenance, l'équipement est remplacé par un équipement équivalent soit à la panne (remplacement correctif) ou lorsque l'âge de l'équipement atteint la valeur T. Une telle politique est dite de type âge.

Dans ce présent travail, l'équipement est supposé fonctionner dans différents environnements. Ces derniers sont supposés avoir un impact sur les performances de l'équipement. Cette performance reflète le dégré de sévérité de l'environnement où l'équipement doit accomplir sa mission. Nous nous limitons au cas où l'équipement est amené à être exploité dans deux environnements distincts de point de vue sévérité. Nous considèrons que la performance de l'équipement est sa fiabilité, la Figure IV.1 montre la performance d'un équipement dans un environnement 1 plus sévère que l'environnement 2.

Dans la suite de cette section, nous supposerons que l'équipement, ayant achevé sa mission dans le premier environnement, est conduit à opérer dans les conditions du second environnement. À la fin de sa première mission, l'âge de l'équipement est évalué à T_1. Ainsi, suite au fait que les conditions opératoires dans le second environnement sont différentes par rapport à celles du premier environnement, l'âge de l'équipement au début de sa seconde mission est donné par une fonction Φ évaluée à T_1 ; c'est à dire $\Phi(T_1)$. Cette fonction de correspondance, liant les deux environnements d'exploitation, tient compte de la sévérité de chaque environnement. De ce fait, si l'environnement 1 est plus sévère que l'environnement 2, alors comme le montre la Figure IV.1, $\Phi(T_1)$ sera supérieur à T_1. Dans le cas contraire, $\Phi(T_1)$ sera inférieure à T_1.

Dans le contexte d'exploitation décrit ci-dessus, l'objectif est de mettre en œuvre une politique optimale de maintenance de type âge. Il s'agit donc de calculer l'âge optimale de remplacement préventif dans les conditions d'exploitation du second environnement.

Dans ce qui suit, nous considérons que le critère d'optimisation est le coût total moyen par unité de temps. Un tel coût est engendré par un coût C_r de remplacement correctif et un coût C_p de remplacement préventif ($C_p << C_r$). Les temps requis pour ces remplacements sont considérés négligeables par rapport à la durée des cycles de fonctionnement de l'équipement.

Dans ce chapitre, nous chercherons à établir le modèle de probabilité du coût moyen unitaire qu'engendre la politique de maintenance, avant de proposer un modèle statistique en utilisant l'estimateur du Noyau de Parzen.

2.1 Modèle de probabilité

Le coût total moyen par unité de temps engendré par les actions de maintenance correctives et préventives est donné par :

$$C\left(T|T_1\right) = \frac{\text{Coût total moyen du remplacement pendant un cycle}}{\text{Durée moyenne du cycle}}, \tag{IV.1}$$

par la suite nous noterons

$$C\left(T|T_1\right) = \frac{C_M}{T_M}. \tag{IV.2}$$

où C_M est le coût total moyen de remplacement lors d'un cycle et T_M la durée moyenne du cycle. Les propositions suivantes permettent de donner les expressions de C_M et T_M. Nous noterons également $R_1(t)$ la fiabilité de l'équipement correspondant à l'environnement d'exploitation 1, et $R_2(t)$ celle correspondant à l'environnement 2. Les durées de vie de l'équipement sont représentées par la variable aléatoire X.

Proposition 3 *Le coût total moyen C_M est donné par*

$$C_M = \frac{1}{R_1(T_1)}\left(C_r R_1(T_1) - (C_r - C_p) R_2\left(T + \Phi(T_1)\right)\right). \tag{IV.3}$$

Preuve. Sachant que l'équipement a T_1 comme âge à la fin de sa mission dans l'environnement 1. Cet âge est évalué par $\Phi(T_1)$ lors du début de sa mission dans le second environnement. Il s'en suit que :

$$C_M = \begin{cases} C_r \text{ si } X \leq T + \Phi(T_1) \\ C_p \text{ si } X > T + \Phi(T_1) \end{cases} \tag{IV.4}$$

d'où

$$C_M = C_r P\left(X \leq T + \Phi(T_1) | X > \Phi(T_1)\right) + C_p P\left(X > T + \Phi(T_1) | X > \Phi(T_1)\right). \tag{IV.5}$$

Avec

$$P\left(X \leq T + \Phi(T_1) | X > \Phi(T_1)\right) = \left(1 - P\left(X > T + \Phi(T_1) | X > \Phi(T_1)\right)\right), \tag{IV.6}$$

nous obtenons :

$$C_M = C_r - (C_r - C_p) P\left(X > T + \Phi(T_1) | X > \Phi(T_1)\right). \tag{IV.7}$$

En utilisant les fonctions de fiabilités dans l'Équation (IV.7), nous obtenons :

$$C_M = C_r - (C_r - C_p) \frac{R_2\left(T + \Phi(T_1)\right)}{R_2\left(\Phi(T_1)\right)}, \tag{IV.8}$$

où $\dfrac{R_2\left(T + \Phi(T_1)\right)}{R_2\left(\Phi(T_1)\right)}$ est la fiabilité conditionnelle de l'équipement dans l'environnement 2 à partir d'un âge $\Phi(T_1)$. Lorsque nous tenons compte de la condition, d'équivalence des fiabilités imposée par la continuité de la fiabilité de l'équipement lors du changement d'environnement.

$$R_2\left(\Phi(T_1)\right) = R_1\left(T_1\right), \tag{IV.9}$$

Alors l'Équation (IV.8) devient,

$$C_M = \frac{1}{R_1\left(T_1\right)} \left(C_r R_1\left(T_1\right) - (C_r - C_p) R_2\left(T + \Phi(T_1)\right)\right). \tag{IV.10}$$

■

De même un cycle se termine à un âge de fonctionnement T ou après une défaillance de l'équipement avant cet âge.

Proposition 4 *Le cycle moyen T_M entre maintenance est donné par :*

$$T_M = \frac{1}{R_1(T_1)} \int_{\Phi(T_1)}^{\Phi(T_1)+T} R_2(t)\mathrm{d}t. \tag{IV.11}$$

Preuve. Rappelons que selon la politique de maintenance adoptée, le cycle se termine soit par une maintenance corrective (occurrence d'une panne avant la date T) ou encore par une

maintenance préventive (l'équipement ayant atteint l'âge T sans panne). Il en résulte que

$$T_M = T_{Mc} + T_{Mp}, \tag{IV.12}$$

où T_{Mc} et T_{Mp} désignent la longueur du cycle lors d'une maintenance corrective et une maintenance préventive et sont définis respectivement par

$$T_{Mc} = \int_0^T uP\left(u + \Phi(T_1) < X \leq u + \Phi(T_1) + du | X > \Phi(T_1)\right), \tag{IV.13}$$

$$T_{Mp} = T * P\left(X > T + \Phi(T_1) | X > \Phi(T_1)\right), \tag{IV.14}$$

$$\begin{aligned} T_M &= T\frac{R_2\left(T + \Phi(T_1)\right)}{R_1(T_1)} + \frac{1}{R_1(T_1)}\int_0^T uf_2\left(u + \Phi(T_1)\right)du, \\ &= T\frac{R_2\left(T + \Phi(T_1)\right)}{R_1(T_1)} - \frac{1}{R_1(T_1)}\int_0^T udR_2\left(u + \Phi(T_1)\right). \end{aligned}$$

Par une intégration par partie, nous obtenons l'expression de la durée moyenne d'un cycle.

$$T_M = \frac{1}{R_1(T_1)}\int_0^T R_2\left(u + \Phi(T_1)\right)\mathrm{d}u. \tag{IV.15}$$

Posons $u + \Phi(T_1) = t$, l'Équation (IV.15) devient,

$$T_M = \int_{\Phi(T_1)}^{\Phi(T_1)+T} \frac{R_2(t)}{R_1(T_1)}\mathrm{d}t. \tag{IV.16}$$

■

D'après les résultats des deux proportions, la fonction objectif à minimiser s'écrit

$$C(T|T_1) = \frac{C_r R_1(T_1) - (C_r - C_p)R_2\left(T + \Phi(T_1)\right)}{\int_{\Phi(T_1)}^{\Phi(T_1)+T} R_2(t)\mathrm{d}t}. \tag{IV.17}$$

Cette fonction vérifie les propriétés suivantes :

$$lim_{T\to\infty}C(T|T_1) = \frac{C_r}{\int_{\Phi(T_1)}^{\infty} R_2(x)\mathrm{d}x}, \tag{IV.18}$$

et

$$lim_{T\to 0}C(T|T_1) = \infty. \tag{IV.19}$$

Les propositions suivantes donnent d'autres propriétés de la fonction répresentant le coût moyen unitaire induit par la maintenance sur le second environnement.

Proposition 5 *Lorsque la distribution des durées de vie de l'équipement est exponentielle alors la fonction coût $C(T|T_1)$ est indépendante de l'age T_1 de fonctionnement dans l'environnement 1.*

Preuve. La distribution des durées de vie de l'équipement étant exponentielle alors la fonction de risque intantanée de défaillance est constante et le processus de vieillissement de l'équipement est exempt de mémoire c'est à dire que pour deux durées t et s positives, nous avons

$$R_2(t+s) = R_2(t)R_2(s), \tag{IV.20}$$

En utilisant l'égalité (IV.20) dans l'Équation (IV.17) nous obtenons :

$$C(T|T_1) = \frac{C_r - (C_r - C_p)R_2(T)}{\int_0^T R_2(x)\mathrm{d}x}. \tag{IV.21}$$

■

Proposition 6 *Lorsque le risque ou taux de défaillance est croissant dans le second environnement alors la fonction objectif définie par l'Équation (IV.17) admet une borne inférieure dont l'expression est définie par la fonction coût déduite de la proposition (5) à travers l'Équation (IV.21).*

Preuve. Pour les fonctions de risque IFR/IHR au delà de $\Phi(T_1)$ alors pour deux réalisations de la variable aléatoire X telles que t et s, alors nous avons :

$$R_2(t+s) \leq R_2(t)R_2(s). \tag{IV.22}$$

De l'Équation (IV.22) nous déduisons l'inégalité des intégrales comme suite

$$\frac{1}{\int_0^T R_2\left(\Phi(T_1)\right) R_2(u)\mathrm{d}u} \leq \frac{1}{\int_0^T R_2\left(\Phi(T_1)+u\right)\mathrm{d}u}. \tag{IV.23}$$

D'un autre côté de l'Équation (IV.22), nous obtenons :

$$C_rR_1(T_1) - (C_r - C_p)R_2\left(\Phi(T_1)\right)R_2(T) \leq C_rR_1(T_1) - (C_r - C_p)R_2\left(\Phi(T_1)+T\right). \tag{IV.24}$$

Tous les termes des inégalités (IV.23) et (IV.24) sont positifs, alors par le produit, les inégalités gardent le même ordre.

$$\frac{C_r R_1(T_1) - (C_r - C_p) R_2\left(\Phi(T_1)\right) R_2(T)}{\int_0^T R_2\left(\Phi(T_1)\right) R_2(u) \mathrm{d}u} \leq \frac{C_r R_1(T_1) - (C_r - C_p) R_2\left(\Phi(T_1) + T\right)}{\int_0^T R_2\left(\Phi(T_1) + u\right) \mathrm{d}u},$$

avec la continuité de la fiabilité lors du changement d'environnements $(R_1(T_1) = R_2\left(\Phi(T_1)\right))$ nous obtenons :

$$\frac{C_r - (C_r - C_p) R_2(T)}{\int_0^T R_2(u) \mathrm{d}u} \leq \frac{C_r R_1(T_1) - (C_r - C_p) R_2\left(\Phi(T_1) + T\right)}{\int_0^T R_2\left(\Phi(T_1) + u\right) \mathrm{d}u}.$$

■

Dans la proposition suivante, nous discuterons de l'existence et l'unicité d'un âge optimal de remplacement pour les équipements à risque intantanée de défaillance $\lambda_2(t)$ croissante.

Proposition 7 *Si $\lambda_2(t)$ est une fonction strictement croissante à partir de l'âge $\Phi(T_1)$ c'est à dire IFR alors il existe un âge optimal T_0 qui garantit un coût de maintenance unitaire minimal. En plus, T_0 est fini si :*

$$\lim_{t \to \infty} \lambda_2(t) > \frac{C_r R_1(T_1)}{(C_r - C_p) \int_{\Phi(T_1)}^{\infty} R_2(u) \mathrm{d}u}. \tag{IV.25}$$

Preuve. Nous savons que $C(T|T_1) = \dfrac{C_M}{T_M}$, par conséquent T_0 est optimal si et seulement si :

$$\left(\frac{dC(T|T_1)}{dT}\right)_{T=T_0} = 0,$$

ce qui conduit à

$$\frac{dC_M}{dT} T_M = \frac{dT_M}{dT} C_M.$$

Des Équations (IV.10) et (IV.16) nous déduisons que,

$$\frac{dC_M}{dT} = -\left(C_r - C_p\right) \frac{1}{R_1(T_1)} \frac{dR_2\left(T + \Phi(T_1)\right)}{dT},$$
$$\frac{dT_M}{dT} = \frac{R_2\left(T + \Phi(T_1)\right)}{R_1(T_1)}.$$

En remplaçant chaque terme par son expression et en divisant l'égalité par $R_2(T+\Phi(T_1))$ nous obtenons :

$$\lambda_2\left(T+\Phi(T_1)\right)\int_0^T R_2\left(u+\Phi(T_1)\right)\mathrm{d}u + R_2\left(T+\Phi(T_1)\right) = \frac{C_r}{C_r-C_p}R_1(T_1). \tag{IV.26}$$

Posons $Lhs(T)$ égale au terme de gauche de l'Équation (IV.26). Alors nous avons une fonction $Lhs(T) = constante$ et avec,

$$Lhs(T) = \lambda_2\left(T+\Phi(T_1)\right)\int_0^T R_2\left(u+\Phi(T_1)\right)\mathrm{d}u + R_2\left(T+\Phi(T_1)\right). \tag{IV.27}$$

La fonction dérivée de $Lhs(T)$,

$$\frac{dLhs(T)}{dT} = \frac{d\lambda_2\left(T+\Phi(T_1)\right)}{dT}\int_0^T R_2\left(u+\Phi(T_1)\right)\mathrm{d}u, \tag{IV.28}$$

puisque $\lambda_2(T)$ est supposé IFR donc $\dfrac{d\lambda_2\left(T+\Phi(T_1)\right)}{dT} > 0$, il s'en suit que la fonction dérivée $\dfrac{dLhs(T)}{dT}$ est aussi positive. Cela implique que la fonction continue $Lhs(T)$ est croissante en fonction de l'âge T. De plus, si nous avons

$$\lim_{T\to\infty} Lhs(T) = \left(\int_{\Phi(T_1)}^{\infty} R_2(u)\mathrm{d}u\right)\lim_{T\to\infty}\lambda_2\left(T+\Phi(T_1)\right) > \frac{C_rR_1(T_)}{C_r-C_p}$$

L'application du théorème des valeurs intermédiaires à $Lhs(T)$ donne l'inégalité $Lhs(0) < \frac{C_rR_1(T_)}{C_r-C_p} < \lim_{T\to\infty} Lhs(T)$. Nous déduisons de cette inégalité l'existence d'une valeur unique $0 < T_0 < \infty$ telle que $Lhs(T_0) = \frac{C_rR_1(T_1)}{C_r-C_p}$ si

$$\lim_{t\to\infty}\lambda_2(t) > \frac{C_rR_1(T_1)}{(C_r-C_p)\int_{\Phi(T_1)}^{\infty} R_2(u)\mathrm{d}u}.$$

Par conséquent pour les équipements dont la fonction de risque instantanée est croissante l'existence d'un âge unique de remplacement optimal dépend de la valeur de la fonction de risque $\lambda_2(T)$ à l'infini. À l'âge de remplacement optimal le risque instantanée de défaillance est défini par :

$$\lambda_2\left(T_0\right) = \frac{C(T_0|T_1)}{C_r-C_p}. \tag{IV.29}$$

■

2.2 Modèle statistique

Rappelons que dans le cadre de nos travaux de thèse, l'équipement considéré est supposé être exploité dans différents environnements ayant un impact direct sur ses performances. En plus nous relaxons l'hypothèse communément adoptée dans la littérature. Selon laquelle la loi des durées de vie de l'équipement est connue. Nous supposons que cette loi doit être estimée à partir des données des durées de vie. Ces données peuvent être hétérogènes. L'objectif est d'établir un modèle statistique de la fonction objectif et d'estimer l'âge optimal de remplacement. Afin d'établir le modèle statistique de la fonction coût de maintenance, remarquons tout d'abord qu'à partir du modèle de probabilité de la fonction objectif, celle-ci peut être complètement estimée par l'estimation des fonctions de fiabilités $R_1(t)$ et $R_2(t)$ de l'équipement correspondants aux conditions d'exploitation respectivement de l'environnement 1 et celles de l'environnement 2. Dans nos travaux, ces fonctions de fiabilité sont estimées par l'exploitation de l'estimateur du Noyau Parzen introduit lors du Chapitre III. Rappelons que, selon cette méthode d'estimation, les fonctions de fiabilité sont données par :

$$\hat{R}_{1h}(T_1) = \frac{1}{n_1}\sum_{i=1}^{n_1} K_s\left(\frac{T_1 - t_{1i}}{h_1}\right), \tag{IV.30}$$

et

$$\hat{R}_{2h}\left(T_2 + \Phi(T_1)\right) = \frac{1}{n_2}\sum_{i=1}^{n_2} K_s\left(\frac{T_2 + \Phi(T_1) - t_{2i}}{h_2}\right), \tag{IV.31}$$

où $\{t_{1i}\}_{i=1,\cdots,n_1}$ et $\{t_{2i}\}_{i=1,\cdots,n_2}$ sont les données des durées de vie relevées respectivement dans l'environnement 1 et 2. La fonction $K_s(t)$ est définit par $K_s(t) = \int_t^\infty K(u)\mathrm{d}u$. À partir de ces équations, le modèle statistique de la fonction objectif est :

$$\hat{C}(T|T_1) = \frac{\dfrac{C_r}{n_1}\displaystyle\sum_{i=1}^{n_1} K_s\left(\frac{T_1 - t_{1i}}{h_1}\right) - \dfrac{C_r - C_p}{n_2}\displaystyle\sum_{i=1}^{n_2} K_s\left(\frac{T + \Phi(T_1) - t_{2i}}{h_2}\right)}{\dfrac{1}{n_2}\displaystyle\int_{\Phi(T_1)}^{\Phi(T_1)+T}\sum_{i=1}^{n_2} K_s\left(\frac{x - t_{2i}}{h_2}\right)\mathrm{d}x} \tag{IV.32}$$

D'après l'équation ci-dessus, le modèle statistique de la fonction coût de maintenance dépend de plusieurs paramètres. Cette fonction dépend naturellement des données des durées de vie ainsi que de leurs tailles, aussi bien de celles relevées lors de l'exploitation de l'équipement dans l'environnement 1 que dans l'environnement 2. Elle dépend également des paramètres de lissage mais également de l'âge T_1 de l'équipement à la fin de sa mission dans le premier environnement. Tous ces paramètres peuvent être exploités pour établir une étude de sensibilité

quant au résultat de l'estimateur de la fonction objectif qui est le coût total par unité de temps $\hat{C}(T|T_1)$. La proposition suivante permet de quantifier l'impact de la variation du paramètre de lissage sur la fonction coût estimée.

Proposition 8 *La fonction coût exprimé par la méthode du Noyau de Parzen (Équation (IV.32)) est insensible aux variations du paramétre de lissage lorsque la fonction noyau satisfait aux conditions émises dans la Proposition (1) à la page (44) et pour une taille des données de durées de vie large.*

Preuve. La preuve de cette proposition est une conséquence des Propositions (1) et (2) du Chapitre III. En effet, la fonction est estimée à partir de l'estimation de la fonction fiabilité de l'équipement. Il en découle donc que la sensibilité de la fonction fiabilité par rapport à la variation du paramètre de lissage est la même que celle de la fonction objectif.

■

L'étude de sensibilité par rapport aux autres paramètres sera établie lors des expérimentations conduites et présentées dans la suite de ce chapitre. Nous discuterons de l'effet de la taille des données ainsi que celui de l'âge T_1 à partir duquel l'équipement sera exploité dans le second environnement.

3 Exemple Numérique

Afin de se rapproche le plus possible des données issues de la réalité, nous considérons, dans le cadre de cet exemple des données de durées de vie générées à partir de mélanges pondérés de lois de probabilités. Dans cet exemple, nous nous limitons à l'exploitation de la distribution de Weibull à deux paramètres. La Table IV.1 donne pour chaque environnement, les valeurs des

Tableau IV.1 – Définition de la distribution source des données

	Environnement 1			Environnement 2		
Paramètre de forme β	1.5	2.0	4.5	2.5	4.0	4.5
Paramètre d'échelle η	10.0			15.0		
Pourcentage de mixture	35%	55%	10%	90%	5%	5%

paramètres des distributions de Weibull ainsi que les poids de pondération. Par exemple dans le cas de l'environnement 1 d'exploitation les données de durées de vie de l'équipement sont supposées de $W(1,5;10)$, $W(2;10)$ et $W(4,5;10)$ avec les proportions respectives 35%, 55% et 10%.

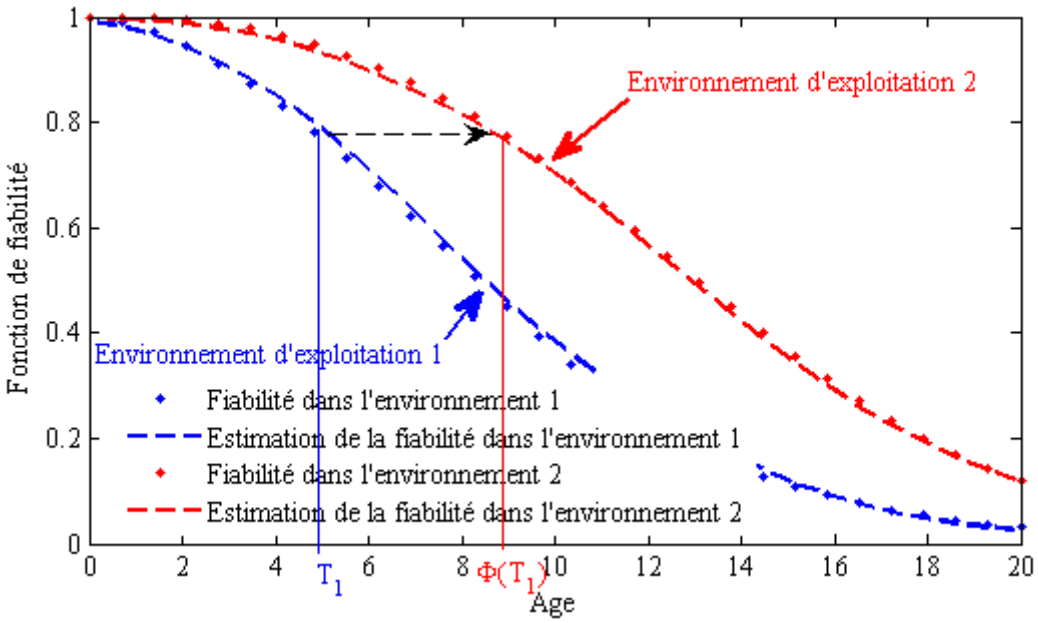

Figure IV.1 – Estimation de la fiabilité par la méthode du plug in de Jones-Sheather J-Sh.

Dans le cadre de cette étude expérimentale, nous avons supposé que les tailles des données n_1 et n_2 collectées respectivement dans l'environnement 1 et l'environnement 2 sont égales ($n_1 = n_2 = n$). La méthode du Noyau de Parzen est exploitée comme méthode d'estimation de la fonction objectif donnée par l'Équation (IV.32). Dans notre travail, deux types d'estimateur du Noyau de Parzen sont implémentés. Le premier estimateur est basé sur la transformation logarithmique des données (identifié par le symbole Log dans ce qui suit). Le deuxième estimateur est basé sur le principe de la méthode du plug-in (noté par J-sh dans la suite du texte). Ces deux estimateurs sont presentés dans le Chapitre III.

A partir des données fournies par la Table IV.1 les échantillons de taille $n = 100, 500, 1000$ des durées de vie sont générées selon les conditions d'exploitation de chaque environnement. La fiabilité estimée dans chaque environnement est représentée par les Figures IV.1 et IV.2. Ces figures sont obtenues pour un échantillon de taille $n = 1000$ de données de durées de vie en utilisant respectivement l'estimateur du Noyau de Parzen avec le plug-in de Jones & Sheater (J-sh) [77] [78] et la transformation logarithmique des données (Log) [43]. D'après les résultats de ces figures, la fiabilité de l'équipement exploité dans l'environnement 1 décroit plus rapidement que dans l'environnement 2. Par conséquent l'environnement 1 est plus sévère que l'environnement 2. De ce fait, un âge T_1 de l'équipement à la fin de sa mission dans le premier environnement correspondrait un âge $\Phi(T_1) > T_1$ au début de sa mission dans le second environnement. Si $T_1 = 5,0$ unités de temps, d'après les résultats de la Figure IV.1,

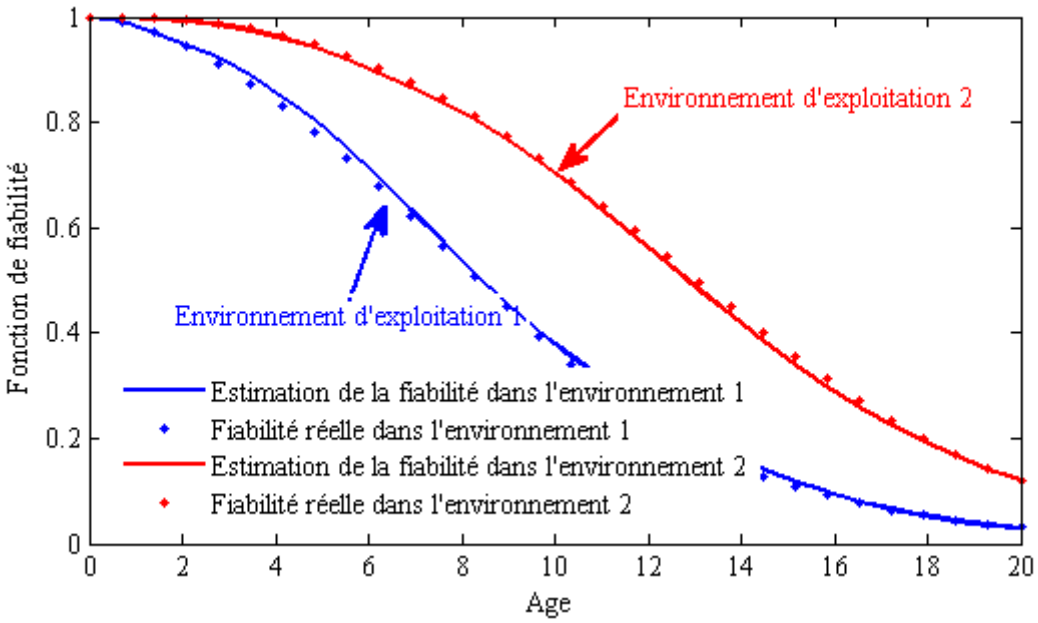

Figure IV.2 – Estimation de la fiabilité par la méthode du difféomorphisme ou Log

l'âge $\Phi(T_1)$ est estimée à $\Phi(T_1) = 9,0$ unités de temps.

La politique de maintenance de type âge étant considérée pour notre équipement. L'équipement est supposé avoir accompli sa mission dans les conditions d'exploitation du premier environnement à la fin duquel l'équipement est d'un âge estimé à T_1 unités de temps. Il est ensuite conduit à accomplir sa mission mais selon les conditions d'exploitation de l'environnement 2 qui, selon les données fixées, est moins sévère que l'environnement 1. La politique de maintenance consiste donc à déterminer l'âge optimal de remplacement préventif dans le nouvel environnement d'exploitation. Pour cela nous supposons que le ratio $\delta_{rp} = \dfrac{C_r}{C_p}$ est fixé à 10. Pour une taille $n = 500$ des données de durées de vie, les Figures IV.3 et IV.4 montrent les variations de la fonction objectif (réel et estimé). Les courbes de la fonction coût réelle et ses estimations sont répresentés par la Figure IV.3, lorsque l'équipement est exploité dans l'environnement 2 après une mission dans l'environnement 1. La Figure IV.4 donne les courbes lorsque l'équipement est exploité dans l'environnement 1 après une mission dans l'environnement 2. Dans le cas de ces figures, l'âge initial T_1 de l'équipement est fixé à $T_1 = 5$.

D'après les Figures IV.3 et IV.4, les deux estimateurs donnent des résultats très proches de la courbe réelle. L'âge optimal de remplacement préventif est évalué approximativement à $T_0 = 3,83$ et $T_0 = 3,72$ respectivement pour le cas des Figures IV.3 et IV.4. Dans le but d'examiner l'effet de la variation de l'âge initial T_1 de l'équipement sur la politique de maintenance, plusieurs

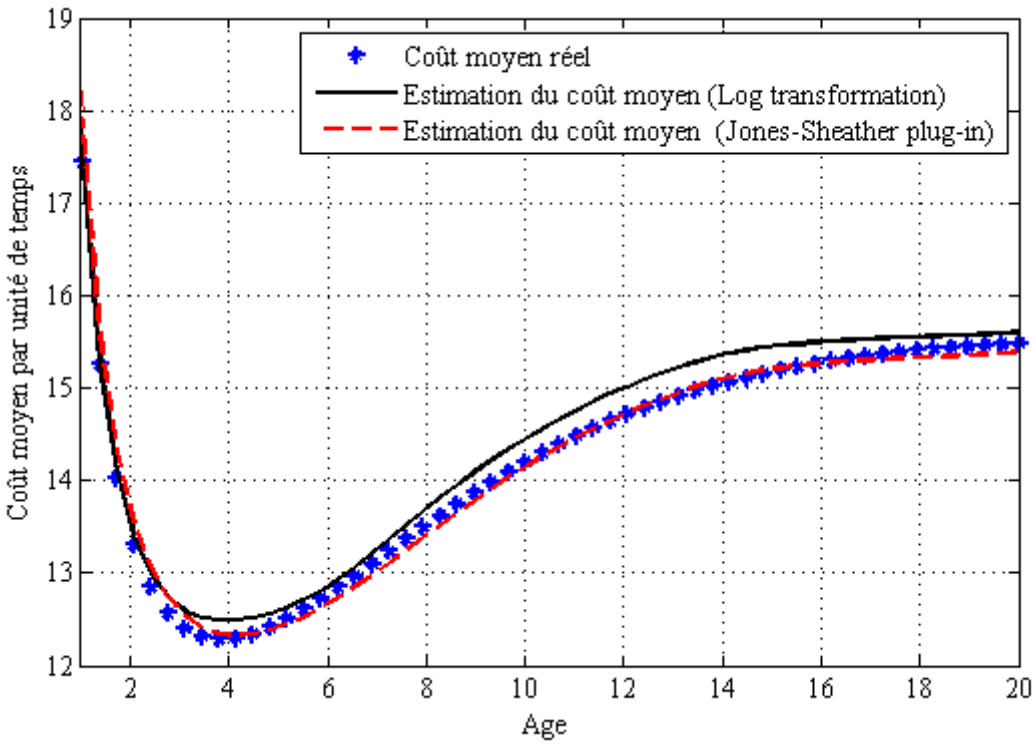

Figure IV.3 – Coût moyen de maintenance par unité de temps par rapport à l'âge : cas où la taille $n = 500$, l'âge initial $T_1 = 5$ et l'équipement est utilisé de l'environnement 1 vers 2.

valeurs de T_1 sont considérées et pour lesquelles le coût de maintenance ainsi que l'âge optimal de remplacement préventif sont estimés. La Figure IV.5 donne les valeurs réelles de T_0 en fonction de T_1. Nous notons que la durée de remplacement optimal T_0 n'est pas monotone avec l'âge initial de l'équipement contrairement aux coûts de la maintenance correspondant Figure IV.6.

Les résultats statistiques du calcul de l'âge optimal sont fournis dans les Tables IV.2-IV.5. Les Tables IV.2 et IV.3 concernent le cas où le système est exploité dans le second environnement après sa mission dans le premier environnement. Les Tables IV.4 et IV.5 concernent le cas inverse où l'équipement est conduit à opérer dans le premier environnement après sa mission dans le second environnement. Afin de tester la robustesse de la méthode d'estimation adoptée, pour chaque valeur de T_1, 10 simulations sont réalisées et ce pour chaque estimateur (Log et J-sh). Ainsi, pour chaque estimation et pour une taille de données de durées de vie. Les Tables IV.2-IV.5 donnent les valeurs moyennes de l'âge optimal de remplacement préventif $E(T_0)$ ainsi que le coût total moyen par unité de temps correspondant. Ces tables donnent également l'écart type *std* ainsi que le coefficient de variation *cv*. Les résultats obtenus démontrent que les estimateurs

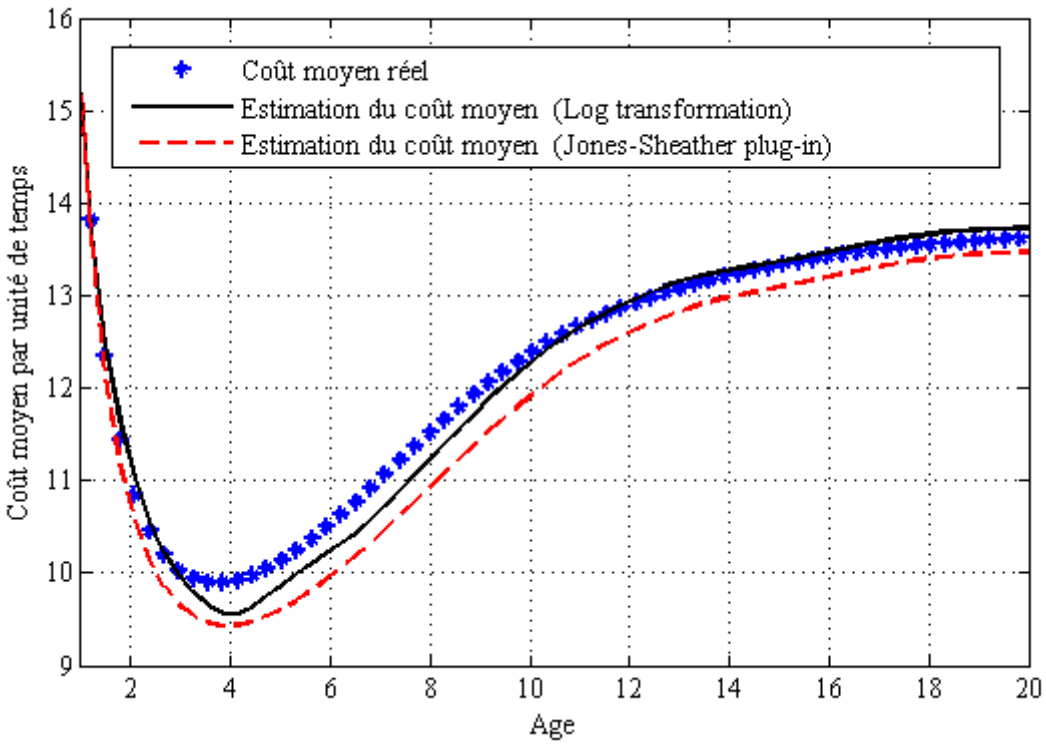

Figure IV.4 – Coût moyen de maintenance par unité de temps par rapport à l'âge : cas où la taille $n = 500$, l'âge initial $T_1 = 5$ et l'équipement est utilisé de l'environnement 2 vers 1.

Log et J-sh sont efficaces même en cas de faibles tailles de données de durées vie. Néanmoins, l'estimateur J-sh comparé à l'estimateur Log fournit de meilleurs résultats. En effet les valeurs de l'écart type (*std*) obtenues dans le cas de l'estimateur J-sh sont inférieures à celles obtenus dans le cas de l'estimateur Log. Ces observations sont confortées par les résultats des Figures IV.7 et IV.8. Pour une taille $n = 1000$ données de durées de vie, les Figures IV.7 et IV.8 représentent, en fonction des différentes valeurs de l'âge initial T_1, les valeurs réelles des dates optimales T_0 ainsi que la moyenne des estimations $E(T_0)$ par l'estimeur J-sh (cas de la Figure IV.7) et celles estimés par l'estimateur Log (cas de la Figure IV.8). D'après les Figures IV.7 et IV.8 toutes les courbes répresentées peuvent être divisée en deux phases. Pendant la première phase allant de 0 jusqu'à l'âge initial $T_1 = 7$, les courbes sont décroissantes. Cela veut dire que dans cette phase, la période de remplacement préventif décroit avec l'accroissement de l'âge initial. En revanche, dans la seconde phase caractérisée par les valeurs de $T_1 > 7$, la période de remplacement préventif croit avec l'accroissement de T_1. En d'autre termes, la politique de maintenance adoptée suggère une période de remplacement préventif qui décroit avec l'âge initial de l'équipement et ce jusqu'à une valeur de $T_1 = 7$ unités de temps. Pour les valeurs

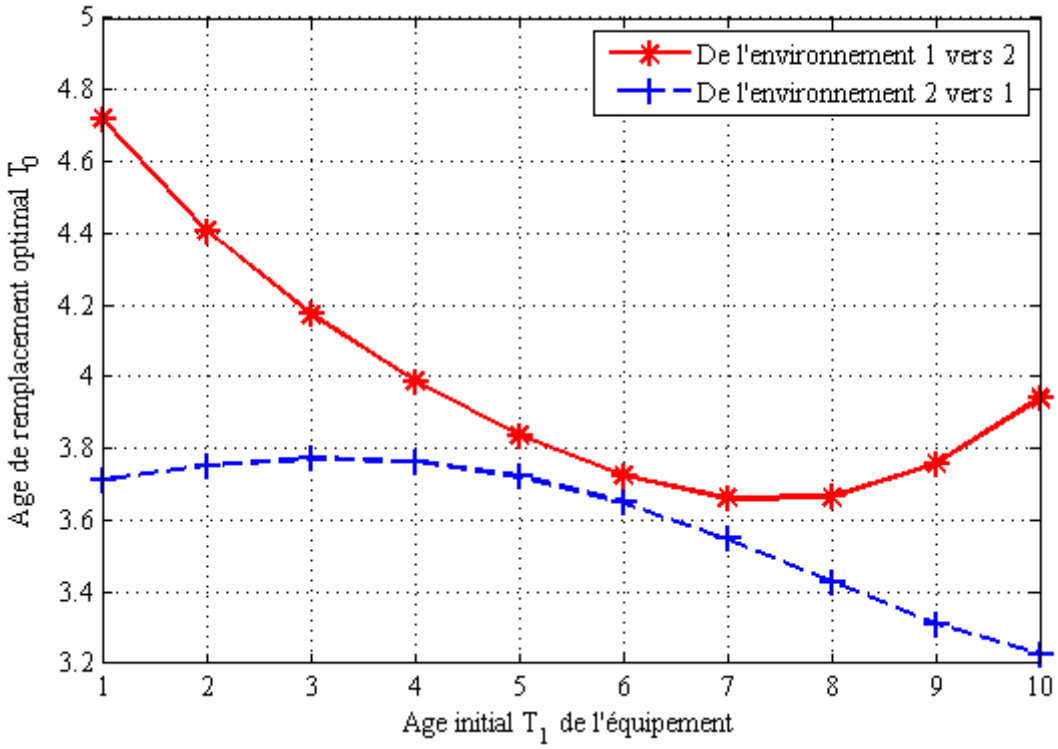

Figure IV.5 – Évolution de l'âge optimal T_0 de remplacement avec l'âge initial T_1.

supérieures à 7 unités de temps, cette politique de maintenance suggère de retarder la date de remplacement à la panne (remplacement correctif). Dans le cas où l'équipement est amené à être exploité dans l'environnement 1 après une période d'exploitation dans l'environnement 2, la politique de maintenance à un effet contraire au cas précedemment analysé Figures IV.9 et IV.10.

Tableau IV.2 – Évaluation de l'âge de remplacement optimal T_0 (de l'environnement 1 à l'environnement 2) $T_1 = 1, 2, 3, 4, 5$

	$n_1 = n_2 = 100$				$n_1 = n_2 = 500$				$n_1 = n_2 = 1000$				—-	
	Log		J-Sh		Log		J-Sh		Log		J-Sh		Real values	
	T_0	Cost	T_0	Cost	T_0	Cost	T_0	Cost	T_0	Cost	T_0	Cost	T_0	Cost
$T_1 = 1$														
E	5.13	6.05	5.17	5.85	4.73	5.52	4.80	5.76	4.62	5.2	4.62	5.47	4.72	5.27
Std	0.53	1.28	0.45	0.62	0.56	0.53	0.33	0.44	0.31	0.44	0.28	0.26		
cv	0.10	0.21	0.09	0.11	0.12	0.1	0.07	0.08	0.07	0.09	0.06	0.05		
$T_1 = 2$														
E	4.62	7.08	4.30	7.08	4.55	7.36	4.62	7.26	4.49	7	4.60	6.96	4.41	6.99
Std	1.01	1.12	1.43	0.91	0.61	0.52	0.25	0.31	0.57	0.35	0.38	0.24		
cv	0.22	0.16	0.33	0.13	0.13	0.07	0.05	0.04	0.13	0.05	0.08	0.03		
$T_1 = 3$														
E	4.04	9	4.09	8.57	4.40	8.69	4.40	8.5	4.20	8.67	4.33	8.53	4.17	8.69
Std	1.06	1.55	0.77	1.09	0.36	0.56	0.15	0.41	0.34	0.75	0.27	0.54		
cv	0.26	0.17	0.19	0.13	0.08	0.06	0.03	0.05	0.08	0.09	0.06	0.06		
$T_1 = 4$														
E	4.31	11.09	4.42	10.73	3.94	10.32	4.02	10.11	4.22	10.24	4.12	10.09	3.99	10.45
Std	0.65	1.63	0.52	1.27	0.58	0.78	0.48	0.58	0.60	0.61	0.35	0.51		
cv	0.15	0.15	0.12	0.12	0.15	0.08	0.12	0.06	0.14	0.06	0.09	0.05		
$T_1 = 5$														
E	4.54	11.26	4.25	10.82	3.76	12.38	3.91	12.05	4.00	12.28	4.00	11.96	3.83	12.28
Std	1.08	1.79	0.77	1.5	0.41	0.84	0.31	0.77	0.24	0.62	0.20	0.48		
cv	0.24	0.16	0.18	0.14	0.11	0.07	0.08	0.06	0.06	0.05	0.05	0.04		

Tableau IV.3 – Évaluation de l'âge de remplacement optimal T_0 (de l'environnement 1 à l'environnement 2) $T_1 = 6, 7, 8, 9, 10$

	$n_1 = n_2 = 100$				$n_1 = n_2 = 500$				$n_1 = n_2 = 1000$				—-	
	Log		J-Sh		Log		J-Sh		Log		J-Sh		Real values	
	T_0	Cost	T_0	Cost	T_0	Cost	T_0	Cost	T_0	Cost	T_0	Cost	T_0	Cost
$T_1 = 6$														
E	5.21	13.89	5.04	13.41	4.42	14.2	4.26	13.8	3.93	14.36	3.91	14.04	3.72	14.22
Std	1.55	2	1.20	1.75	0.65	0.99	0.49	0.96	0.41	0.35	0.31	0.3		
cv	0.30	0.14	0.24	0.13	0.15	0.07	0.12	0.07	0.10	0.02	0.08	0.02		
$T_1 = 7$														
E	3.70	16.1	4.24	16.04	4.08	15.85	3.80	15.55	3.82	16.78	3.77	16.24	3.66	16.25
Std	0.85	1.47	1.92	1.9	0.76	0.63	0.49	0.56	0.43	1.01	0.19	0.63		
cv	0.23	0.09	0.45	0.12	0.19	0.04	0.13	0.04	0.11	0.06	0.05	0.04		
$T_1 = 8$														
E	7.73	18.01	3.99	19.42	3.25	18.23	3.73	17.83	4.15	17.62	3.85	17.52	3.66	18.38
Std	4.42	3.23	0.62	7	1.14	1.2	0.39	0.87	0.57	0.74	0.36	0.74		
cv	0.57	0.18	0.16	0.36	0.35	0.07	0.10	0.05	0.14	0.04	0.09	0.04		
$T_1 = 9$														
E	5.89	20.15	4.52	20.04	4.22	20.04	3.83	19.93	4.03	20.44	4.20	20.37	3.76	20.56
Std	2.56	2.2	2.97	2.19	0.33	1.28	0.28	1.19	1.36	0.75	0.43	0.72		
cv	0.44	0.11	0.66	0.11	0.08	0.06	0.07	0.06	0.34	0.04	0.10	0.04		
$T_1 = 10$														
E	4.80	20.15	3.70	20.77	4.85	22.17	4.25	22.54	4.65	21.53	4.56	21.66	3.94	22.69
Std	2.39	2.05	1.48	2.47	0.93	1.53	0.52	1.2	0.84	1.04	0.80	0.87		
cv	0.50	0.1	0.40	0.12	0.19	0.07	0.12	0.05	0.18	0.05	0.18	0.04		

Tableau IV.4 – Évaluation de l'âge de remplacement optimal T_0 (de l'environnement 2 à l'environnement 1) $T_1 = 1, 2, 3, 4, 5$

	$n_1 = n_2 = 100$				$n_1 = n_2 = 500$				$n_1 = n_2 = 1000$				—	
	Log		J-Sh		Log		J-Sh		Log		J-Sh		Real values	
	T_0	Cost	T_0	Cost	T_0	Cost	T_0	Cost	T_0	Cost	T_0	Cost	T_0	Cost
$T_1 = 1$														
E	3.66	9.72	4.08	7.35	3.5	8.46	3.9	6.47	3.63	8.75	3.68	6.63	3.71	6.82
Std	0.71	1.22	0.59	0.73	0.61	0.47	0.46	0.31	0.45	0.39	0.17	0.25		
cv	0.19	0.13	0.15	0.1	0.17	0.06	0.12	0.05	0.12	0.04	0.05	0.04		
$T_1 = 2$														
E	3.88	9.44	3.99	7.38	3.78	8.46	3.93	6.92	3.8	8.5	3.8	6.74	3.75	7.4
Std	1.17	1.41	0.59	0.98	0.7	0.86	0.44	0.56	0.35	0.31	0.15	0.36		
cv	0.3	0.15	0.15	0.13	0.19	0.1	0.11	0.08	0.09	0.04	0.04	0.05		
$T_1 = 3$														
E	3.27	8.93	3.82	7.26	3.82	8.63	3.84	7.82	3.7	8.65	3.76	7.94	3.77	8.13
Std	1.09	1.31	0.43	1.15	0.37	0.58	0.19	0.43	0.43	0.45	0.16	0.4		
cv	0.33	0.15	0.11	0.16	0.1	0.07	0.05	0.05	0.12	0.05	0.04	0.05		
$T_1 = 4$														
E	3.45	9.78	3.79	8.03	3.81	9.43	3.75	8.59	3.57	9.12	3.7	8.82	3.76	8.96
Std	1.49	1.74	0.37	1.22	0.6	1.33	0.36	0.36	0.24	0.51	0.15	0.43		
cv	0.43	0.18	0.1	0.15	0.16	0.14	0.09	0.04	0.07	0.06	0.04	0.05		
$T_1 = 5$														
E	3.87	10.72	3.96	9.9	3.74	9.87	3.81	9.51	3.68	9.8	3.67	9.62	3.72	9.89
Std	0.84	1.58	0.64	1.4	0.59	0.65	0.35	0.54	0.26	0.53	0.18	0.44		
cv	0.22	0.15	0.16	0.14	0.16	0.07	0.09	0.06	0.07	0.05	0.05	0.05		

Tableau IV.5 – Évaluation de l'âge de remplacement optimal T_0 (de l'environnement 2 à l'environnement 1) $T_1 = 6, 7, 8, 9, 10$

	$n_1 = n_2 = 100$				$n_1 = n_2 = 500$				$n_1 = n_2 = 1000$				—	
	Log		J-Sh		Log		J-Sh		Log		J-Sh		Real values	
	T_0	Cost	T_0	Cost	T_0	Cost	T_0	Cost	T_0	Cost	T_0	Cost	T_0	Cost
$T_1 = 6$														
E	3.96	10.63	3.84	10.34	3.83	11.22	3.61	10.82	3.74	10.95	3.7	10.83	3.65	10.94
Std	0.84	1.03	0.71	0.86	0.73	0.89	0.41	0.66	0.46	0.62	0.35	0.52		
cv	0.21	0.1	0.18	0.08	0.19	0.08	0.11	0.06	0.12	0.06	0.09	0.05		
$T_1 = 7$														
E	4.53	12.64	3.96	11.77	3.68	11.78	3.66	11.68	3.44	12.37	3.63	12.17	3.55	12.1
Std	1.3	1.69	0.69	0.89	0.4	0.78	0.28	0.53	0.86	0.47	0.24	0.51		
cv	0.29	0.13	0.17	0.08	0.11	0.07	0.08	0.05	0.25	0.04	0.07	0.04		
$T_1 = 8$														
E	4.43	13.23	3.45	12.65	3.52	14.35	3.81	13.73	3.45	13.73	3.44	13.5	3.43	13.41
Std	1.77	1.06	0.47	1.06	0.58	0.8	0.81	0.75	0.36	0.59	0.25	0.5		
cv	0.4	0.08	0.14	0.08	0.17	0.06	0.21	0.05	0.1	0.04	0.07	0.04		
$T_1 = 9$														
E	4.3	15.14	3.67	14.32	4.09	15.12	3.57	14.72	3.36	15.12	3.36	14.93	3.31	14.89
Std	1.67	1.37	0.76	0.81	1.42	1.71	0.29	1.01	0.39	0.71	0.28	0.64		
cv	0.39	0.09	0.21	0.06	0.35	0.11	0.08	0.07	0.12	0.05	0.08	0.04		
$T_1 = 10$														
E	4.79	15.84	4.07	15.33	3.68	16.83	3.45	16.5	3.25	16.8	3.25	16.53	3.23	16.55
Std	1.4	1.67	0.81	1.74	0.72	1.59	0.52	1.26	0.66	0.85	0.5	0.73		
cv	0.29	0.11	0.2	0.11	0.19	0.09	0.15	0.08	0.2	0.05	0.15	0.04		

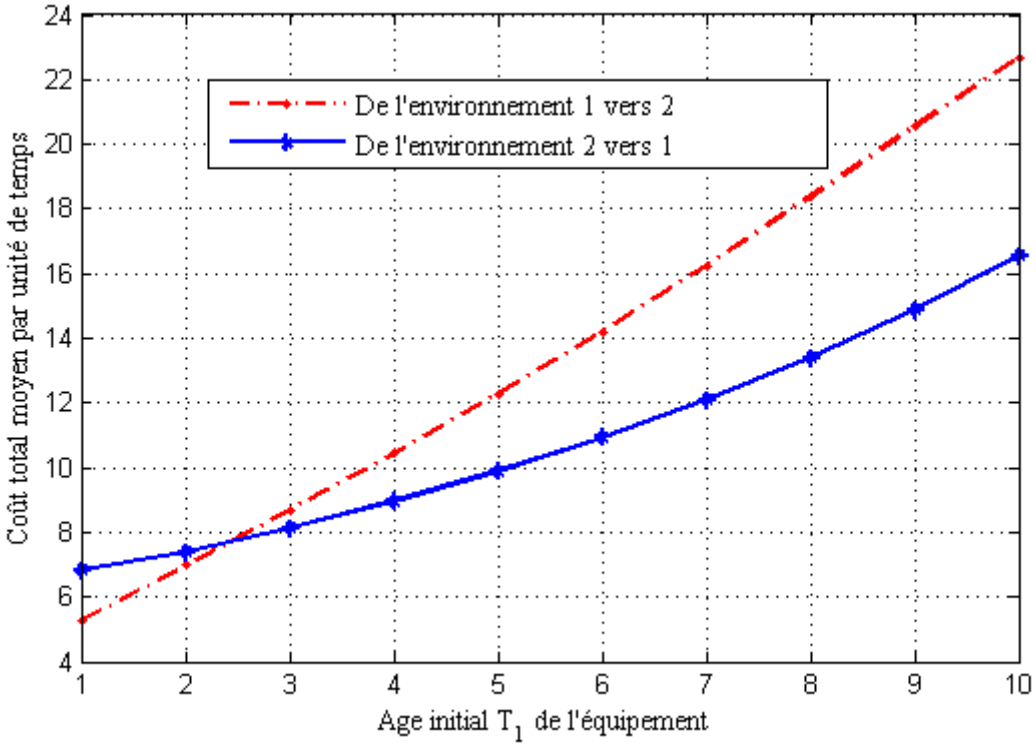

Figure IV.6 – Évolution de coût optimal C_0 en fonction de l'âge initial T_1.

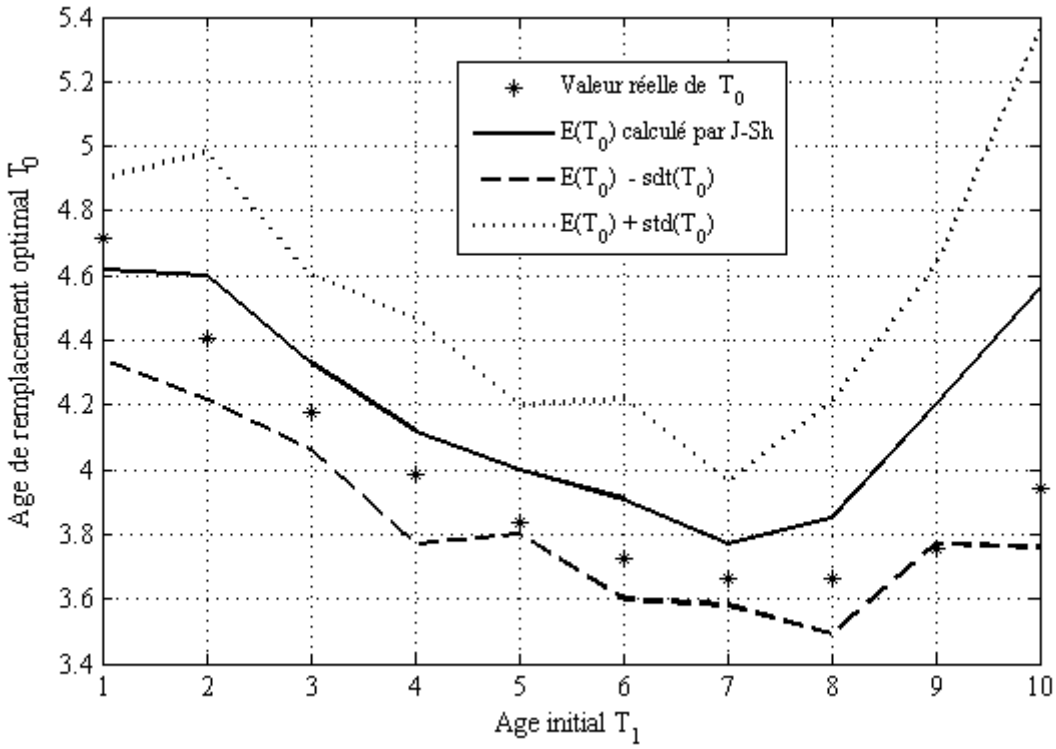

Figure IV.7 – Intervalle de variation de l'âge optimal à un écart type près, cas de l'estimation J-Sh (environnement 1 vers 2).

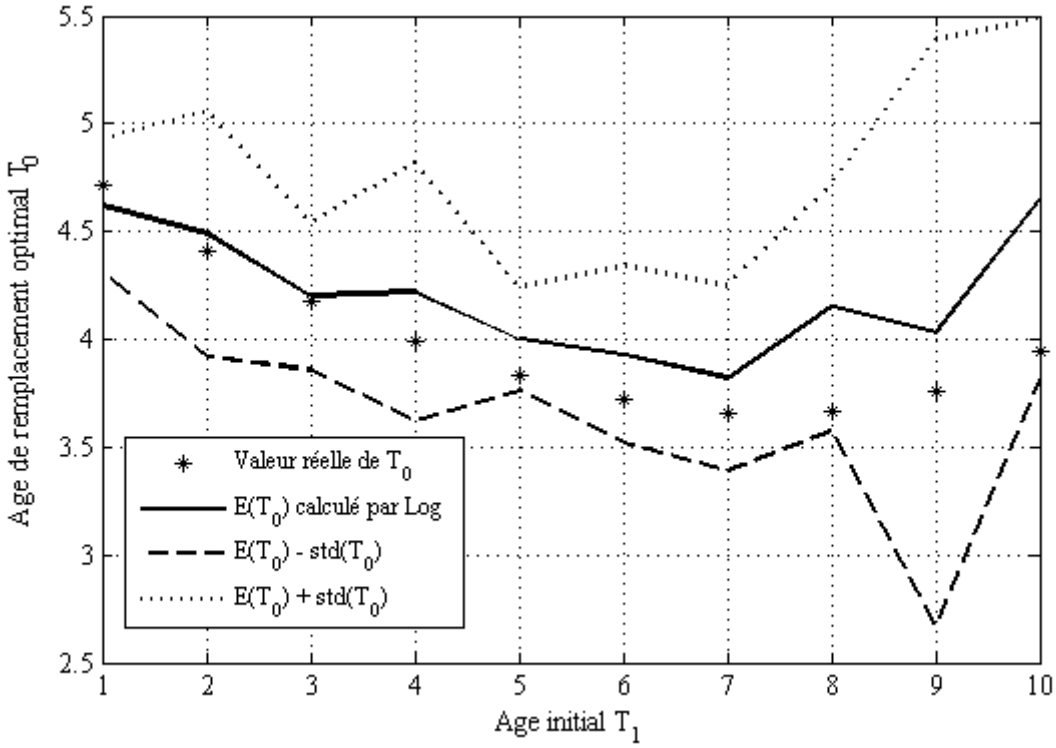

Figure IV.8 – Intervalle de variation de l'âge optimal à un écart type près, cas de l'estimation Log (environnement 1 vers 2).

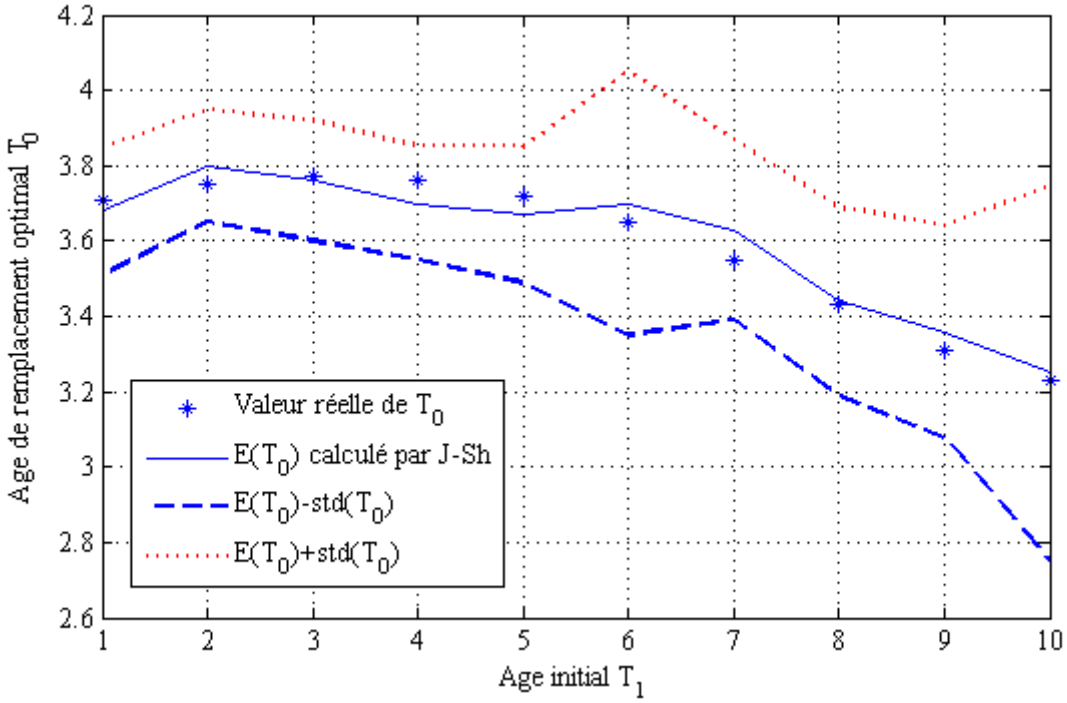

Figure IV.9 – Intervalle de variation de l'âge optimal à un écart type près, cas de l'estimation J-Sh (environnement 2 vers 1).

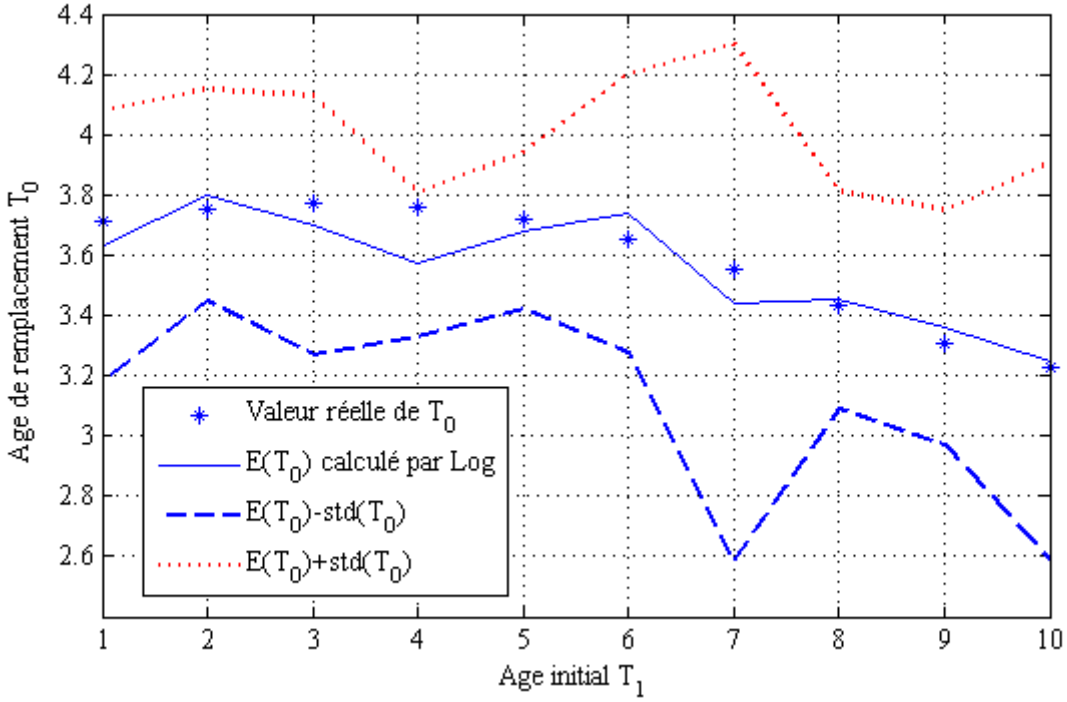

Figure IV.10 – Intervalle de variation de l'âge optimal à un écart type près, cas de l'estimation Log (environnement 2 vers 1).

4 Politique de maintenance de type âge avec réparation

La section précedente a traité l'optimisation de la maintenance des équipements ayant un certain âge de fonctionnement dans un environnement distinct. Dans cette formulation, seul les coûts induits par la maintenance dans le second environnement sont pris en compte. L'historique du coût de maintenance dans le premier environnement est ignoré et l'équipement est supposé ne subir aucune remise à niveau avant le début de l'opération sur le second environnement.

L'analyse théorique de cette section prend en compte les coûts de réparations réalisées dans le premier environnement dans l'optimisation de la maintenance. Les nouvelles politiques seront théoriquement analysée dans les sections suivantes, où les conditions sur l'unicité et l'existence des solutions optimales seront discutées.

4.1 Modèle mathématique avec réparation à T_1

Dans cette nouvelle politique, nous maintenons les mêmes conditions d'exploitation de l'équipement ainsi que l'hypothèse sur l'horizon d'exploitation (infini). Ce qui nous permet de garder la même formulation de la fonction coût comme décrite dans l'Équation (IV.17). Ainsi, à ce modèle de coût nous ajoutons le coût supplémentaire rélativement à la réparation effectuée à l'âge T_1 juste avant le début de l'opération dans le second environnement. Cette réparation a pour objectif de réduire l'âge de l'équipement de T_1 à t_1 et de lui permettre de débuter les opérations dans le second environnement avec un niveau de fiabilité meilleure.

Le coût de la réparation permettant la réduction d'âge de T_1 à t_1 à la fin de l'opération dans le premier environnement est noté par $C_{RM}(t_1|T_1)$ et défini sous les conditions aux limites suivantes.

$$\lim_{t_1 \to 0^+} C_{RM}(t_1|T_1) = \infty, \tag{IV.33}$$

$$\lim_{t_1 \to T_1^-} C_{RM}(t_1|T_1) = 0. \tag{IV.34}$$

La condition (IV.33) traduit l'absence d'une réparation garantissant une réduction complète de l'âge à partir de T_1 (c'est à dire débuter avec $t_1 = 0$). Tandis que la condition (IV.34) montre que la non-réparation de l'équipement à T_1 n'engendre pas de coût. En plus la fonction coût $C_{RM}(t_1|T_1)$ doit être monotone en fonction des variables t_1 et T_1. Ses fonctions dérivées

rélativement aux variables t_1 et T_1 doivent être respectivement négative et positive :

$$\frac{d}{dt_1} C_{RM}(t_1|T_1) < 0, \tag{IV.35}$$

$$\frac{d}{dT_1} C_{RM}(t_1|T_1) > 0. \tag{IV.36}$$

En se basant sur l'Équation (IV.17 de la page 85), le coût total moyen par unité de temps s'écrit alors :

$$C\left(T|t_1, T_1\right) = \frac{\left(C_{RM}(t_1|T_1) + C_r\right) R_1\left(t_1\right) - \left(C_r - C_p\right) R_2\left(T + \Phi(t_1)\right)}{\int_0^T R_2\left(\Phi(t_1) + u\right) \mathrm{d}u}, \tag{IV.37}$$

où l'on vérifie les propriétés suivantes :

$$\lim_{T \to 0} C\left(T|t_1, T_1\right) = \infty, \tag{IV.38}$$

$$\lim_{T \to \infty} C\left(T|t_1, T_1\right) = \frac{C_{RM}(t_1|T_1) + C_r}{\int_0^\infty R_2\left(\Phi(t_1) + u\right) \mathrm{d}u} R_1(t_1). \tag{IV.39}$$

L'Équation (IV.37) admet une solution optimale pour un couple d'âge (t_1, T_1) si la dérivée du coût total moyen unitaire $C(T|t_1, T_1)$ par rapport T s'annule :

$$\frac{d}{dT} C(T|t_1, T_1) = 0. \tag{IV.40}$$

Cette équation est équivalente à :

$$Lhs(T|t_1) = G(t_1, T_1), \tag{IV.41}$$

où

$$Lhs(T|t_1) = \lambda_2\left(T + \Phi(t_1)\right) \int_0^T R_2\left(\Phi(t_1) + u\right) \mathrm{d}u + R_2\left(\Phi(t_1) + T\right),$$

et

$$G(t_1, T_1) = \frac{C_{RM}(t_1|T_1) + C_r}{C_r - C_p} R_1(t_1).$$

La proposition suivante donne les conditions nécessaires et suffisantes pour l'existence et l'unicité d'un âge optimal de remplacement dans le second environnement.

Proposition 9 *Il existe une stratégie optimale de remplacement préventif (dans le second environnement d'exploitation) si le taux de panne $\lambda_2(t)$ est croissant et vérifie l'une des deux conditions :*

1. $\lim_{T\to\infty} \lambda_2(T) = \infty$.
2. $\lim_{T\to\infty} \lambda_2(T) > \dfrac{C_{RM}(t_1|T_1) + C_r}{(C_r - C_p)\int_0^\infty R_2(t)dt} R_1(t_1)$.

Preuve. Nous savons que le remplacement préventif optimal à un âge pendant l'exploitation serait réalisable si l'Équation (IV.40) admet une solution finie. Alors pour un couple d'âges (t_1, T_1), nous déduisont de l'Équation (IV.40) la condition suivante :

$$Lhs(T|t_1) = G(t_1, T_1),$$

où $Lhs(T|t_1)$ est une fonction continue et $G(t_1, T_1)$ une constante rélativement à T. La fonction dérivée de $Lhs(T|t_1)$ par rapport à T donne :

$$\frac{d}{dT} Lhs(T|t_1) = \left(\int_0^\infty R_2\left(\Phi(t_1) + u\right) \mathrm{d}u\right) \frac{d}{dT}\lambda_2(T) > 0.$$

Puisque le taux de panne est croissant $\frac{d}{dT}\lambda_2(T) > 0$, alors la fonction $\frac{d}{dT}Lhs(T|t_1) > 0$ et la fonction $Lhs(T|t_1)$ est par conséquent croissante. Dans ce cas, pour garantir une solution, il suffit que :

$$\lim_{T\to\infty} Lhs(T|t_1) > G(t_1, T_1),$$

en d'autres termes,

$$\lim_{T\to\infty} \lambda_2(T) \int_0^\infty R_2\left(\Phi(t_1) + u\right) \mathrm{d}u + R_2(\infty) > G(t_1, T_1),$$

où l'on peut donc déduire,

$$\lim_{T\to\infty} \lambda_2(T) > \frac{C_{RM}(t_1|T_1) + C_r}{(C_r - C_p)\int_0^\infty R_2(t)dt} R_1(t_1).$$

.

■

Nous remarquons que la condition 1 de la Proposition (9) est relative aux équipements dont le taux de panne est strictement croissant. Pour de tels équipements, il existe toujours une stratégie de remplacement optimale dans le second environnement quels que soient le coût et la réduction d'âge induits par la réparation avant usage. Pour de tels équipements, le taux de panne $\lambda_2(t)$ est croissant et non-borné. Par contre, lorsque la distribution des durées de vie admet un taux de panne croissant mais borné (cas des distributions Gamma (paramètre de forme $\beta > 1$) et une Log-Normale), alors la stratégie optimale est déterminée par la condition

2 de la Proposition (9). Dans ce cas de figure, l'optimalité dépend de la limite à l'infini ou de la valeur maximale du taux de panne. Cependant, si la limite à l'infini ou la valeur maximale du taux de panne est inférieur à $\dfrac{C_{RM}(t_1|T_1)+C_r}{(C_r-C_p)\int_0^\infty R_2(t)dt}R_1(t_1)$ alors il n'y a pas de stratégie de remplacement optimale et l'équipement pourra être utilisé jusqu'à la défaillance parce que l'âge optimal de remplacement sera égal à $T_0=\infty$.

4.2 Modèle avec réparation minimale avant T_1

Jusque là les coûts de réparation induit par l'exploitation dans le premier environnement ont été négligés dans la modélisation. Afin d'atteindre une modélisation sur toute l'étendue de l'exploitation, le coût de maintenance qu'engendre l'exploitation dans le premier environnement est pris en compte et noté par $C_m(t)$. Ce coût résulte de la réparation minimale que subira l'équipement pendant le premier environnement en vue de le maintenir au même état qu'avant la défaillance. C'est à dire que le processus d'exploitation s'articule au tour de deux environnements d'exploitation. Sur le premier environnement l'équipement subit une réparation minimale et est remplacé sur le second environnement après défaillance. Cette dernière politique de maintenance se rapproche du concept du Burn-in intégrée avec l'optimisation des politiques de maintenance. Le Burn-in consiste à protéger le client des défaillances précoses pour les équipements ayant une fonction risque instantané en baignoire (présence d'un intervalle de rodage sur lequel la fonction de risque instantané décroit). Ce concept est expliqué dans le livre de Jensen et Petersen [93] et intégré avec les politiques de maintenance dans les travaux de Jie-Mi [94] et Cha [95], [96]. Notre modèle est mathématiquement plus proche de celui de Cha que celui de Jie Mi. Ainsi que ces auteurs, notre objectif demeure la recherche d'un âge optimal de maintenance garantissant un coût de maintenance minimal à l'exploitation. La fonction coût totale résulte de l'intégration des coûts liés à la réparation minimale dans le premier environnement sur l'intervalle $[0,T_1]$ au coût de la politique de maintenance de type âge dans le second environnement. Le coût de la réparation minimale est définie comme suite :

$$C_m(T_1) = C_m\Lambda_1(T_1), \tag{IV.42}$$

$$= C_m\int_0^{T_1}\lambda_1(t)\mathrm{d}t, \tag{IV.43}$$

$$= -C_m\log\left(R_1(T_1)\right). \tag{IV.44}$$

avec C_m le coût d'une répartion minimale et $\Lambda_1(T_1)$ le risque de défaillance cumulé qui est égal au nombre moyenne de défaillance sur l'intervalle $[0,T_1]$ sous une réparation minimale. La durée

moyenne d'un cycle dans le second environnement est augmenté de l'âge T_1. Alors le coût total moyen de la maintenance par unité de temps est :

$$C(T|T_1) = \frac{C_m(T_1) + \frac{1}{R_1(T_1)}\left(C_r R_1(T_1) - (C_r - C_p) R_2\left(\Phi(T_1) + T\right)\right)}{T_1 + \frac{1}{R_1(T_1)} \int_0^T R_2\left(\Phi(T_1) + u\right) \mathrm{d}u}, \tag{IV.45}$$

$$= \frac{\left(C_m(T_1) + C_r\right) R_1(T_1) - (C_r - C_p) R_2\left(\Phi(T_1) + t\right)}{T_1 R_1(T_1) + \int_0^T R_2\left(\Phi(T_1) + u\right) \mathrm{d}u}, \tag{IV.46}$$

$$\tag{IV.47}$$

avec

$$lim_{T\to 0} C(T|T_1) = \frac{C_m(T_1) + C_p}{T_1}, \tag{IV.48}$$

$$lim_{T\to\infty} C(T|T_1) = \frac{C_m(T_1) + C_r}{T_1 R_1(T_1) + \int_0^\infty R_2\left(\Phi(T_1) + u\right) \mathrm{d}u} R_1(T_1). \tag{IV.49}$$

Tout comme les modèles ci-dessus, il existe un âge optimal de remplacement dans le second environnement si :

$$\frac{d}{dT} C(T|T_1) = 0, \tag{IV.50}$$

cette équation conduit à

$$Lhs(T) = G(T_1), \tag{IV.51}$$

avec

$$Lhs(T) = \lambda_2\left(\Phi(T_1) + T\right)\left(\int_0^T R_2\left(\Phi(T_1) + u\right) \mathrm{d}u + T_1 R_1(T_1)\right) + R_2\left(\Phi(T_1) + T\right),$$

et

$$G(T_1) = \frac{C_r + C_m \Lambda_1(T_1)}{C_r - C_p} R_1(T_1),$$

Proposition 10 *Pour un équipement à taux de panne croissant (IFR/IHR), il existe alors une solution unique si :*

$$\lambda_1\left(\Phi(T_1)\right) T_1 < \frac{C_p + C_m \Lambda_1(T_1)}{C_r - C_p}, \tag{IV.52}$$

et

$$\lim_{t\to\infty} \lambda_2(t) > \frac{C_p + C_m \Lambda_1(T_1)}{C_r - C_p} \frac{R_1(T_1)}{T_1 R_(T_1) + \int_0^\infty R_2\left(\Phi(T_1) + u\right) \mathrm{d}u}, \tag{IV.53}$$

en plus cette unicité s'étend à T_1 si $C_r > C_m$ (cette preuve est triviale et s'explique à partir de la monotonie de $G(T_1)$ par rapport T_1).

Preuve. La démonstration est identique à celle de la Proposition (9) de la section précédente. ■

5 Conclusion

Dans ce chapitre, nous avons proposé un modèle statistique d'optimisation des politiques de maintenance appliquées à des équipements usagés. Dans cette étude, l'horizon d'optimisation des politiques de maintenance est supposé infini. Les lois de distribution des durées de vie des équipements ont été supposées a priori inconnues. Les équipements, soumis à ces politiques de maintenance, ne sont caractérisés que par des échantillons de données de durées de vie indépendantes. Les mêmes équipements sont ensuite considérés évoluer succéssivement dans plusieurs environnements ou sous plusieurs conditions d'exploitation distincts. Le premier objectif de ce chapitre a été la modélisation et l'analyse des coûts moyens unitaires des différentes politiques de maintenance. Dans ce chapitre, les fonctions coût des politiques de maintenance sont établies et discutées afin d'élucider les conditions d'optimalité de chacune des stratégies adoptées. Nous avons montré, pour les équipements dont la fonction risque instantané de défaillance est croissante et non-bornée, qu'il existe toujours un âge optimal de maintenance préventive. Cet âge de maintenance garantit un coût d'exploitation minimal de l'équipement. Le deuxième objectif majeur de ce chapitre a été l'estimation de cet âge optimal car les fonctions de distribution des durées de vie des équipements ont été supposées inconnues. Les fonctions coût des politiques de maintenance sont évaluées à travers l'estimation des fonctions de distribution des durées de vie par la méthode du Noyau de Parzen. La méthode du Noyau de Parzen est un estimateur non-paramétrique dont l'implémentation nécessite une fonction noyau de densité de probabilité et de paramètre de lissage. Le paramètre de lissage est déterminant dans la convergence de l'estimateur du Noyau de Parzen. Ainsi, nous avons prouvé dans ce chapitre que la variabilité et l'effet de l'erreur dans le calcul du paramètre de lissage à une influence faible sur l'évaluation et l'optimisation de la fonction coût. Par la suite, un exemple numérique est proposé et dans lequel, les durées de vie des équipements ont éte générées de manière aléatoire. Nous avons évalué l'âge optimal de maintenance préventive et analysé sa sensibilité par rapport à l'âge initial de l'équipement. Nous avons également montré que le coefficient de variation de l'estimateur de l'âge optimal en utilisant la méthode du Noyau de Parzen est faible et diminue avec la taille des données de durées de vie. L'analyse faite dans cet exemple numérique a également

montré que les estimations diffèrent très peu des valeurs réelles de l'âge optimal calculées à partir du modèle théorique. Les résultats obtenus et exposés dans ce chapitre ont fait l'objet des publications dans [97] [98].

Ce chapitre a aussi permis l'étude de quelques extensions de la politique de maintenance décrite dans le paragraphe précédent. Dans l'une de ces extensions, l'équipement usagé subit initialement une réparation pour atteindre un niveau de fiabilité requis avant son exploitation. Dans la seconde extension étudiée, nous avons intégré les coûts de maintenance antérieurs de l'équipement. En supposant que les maintenances antérieures subies par l'équipement sont des réparations minimales. Ces extensions ont été étudiées et nous avons montré l'existence d'un âge de maintenance préventive optimal pour les équipements dont le risque de défaillance instantané est croissant et non-borné. Cependant, la mise en place opérationnelle de telles stratégies de maintenance exige une exécution des opérations de maintenance préventive à la date optimale préalablement établie et sans retard. Cette contrainte opérationnelle est difficile à satisfaire en réalité à cause de la disponibilité des pièces de réchange ou de l'équipe de maintenance, de la demande du marché ou encore les conditions environnementales. Le Chapitre V proposera une approche permettant de mesurer et d'analyser les conséquences du report de la maintenance préventive.

Chapitre V

Analyse de la maintenance préventive retardée

Le Chapitre V proposera une approche permettant de mesurer et d'analyser le risque induit par le report d'une maintenance préventive. Ce report peut être dû à l'exploitation d'équipement telles qu'un pic de demande, un manque de pièces de rechange, une indisponibilité des intervenants ou une perturbation météorologique. Nous proposerons une analyse du risque relatif au report de la maintenance préventive indépendamment de la politique de maintenance adoptée. Cette analyse repose sur une fonction de risque proposée.

1 Introduction

La majorité des travaux existant dans la littérature et qui traitent de l'optimisation des stratégie de maintenance sont basées sur une hypothèse qui stipule que les actions de maintenance préventive sont exécutées sans aucun retard. Les modèles mathématiques élaborés dans ces travaux permettent en effet le calcul de la date optimale de maintenance préventive. Parmi les hypothèses communément adoptées, nous pouvons citer par exemple la disponibilité des pièces de rechange et celle des opérateurs en maintenance. En pratique, il existe bien des cas où ces hypothèses ne peuvent être toutes ou en partie satisfaites. Une hypothèse non satisfaite entraîne un retard quant à la date initialement prévue pour les actions de maintenance préventive.

La rupture des stocks des pièces de réchange et l'indisponibilité des opérateurs en maintenance sont des contraintes internes à l'entreprise qui peuvent induire des délais pour les actions de maintenance préventives. D'autres contraintes externes à l'entreprise peuvent également avoir les mêmes conséquences que les contraintes internes. À titre d'exemple, les conditions environnementales d'exploitation de l'équipement à maintenir peuvent retarder l'exécution de la maintenance notamment pour les équipements offshores. Pour de tels équipements les conditions météorologiques peuvent entraîner des risques humains et matériels lors de l'accomplissement des activités de maintenance. Ces activités peuvent être affectées par des conditions liées à la satisfaction des demandes du marché. Dans ce dernier cas de figure, les actions de maintenance sont retardées le temps nécessaire de répondre complètement ou partiellement à la demande du marché. Retarder les activités de maintenance peut avoir des inconvénients mais également des avantages. En effet quand il s'agit de vie humaine en plus de l'endommagement de matériels, il est évident que retarder la maintenance est le choix raisonnable. Aussi retarder la maintenance des équipements pour faire face à une demande du marché peut engendrer beaucoup de bénéfice pour l'entreprise. Dans tous les cas de figure, l'inconvénient résultant de tels retards est l'accroissement de détérioration de l'équipement et la probabilité croissante du risque de sa défaillance. Un trad-off existe donc entre exploiter l'équipement sans maintenance et diminuer le risque de sa défaillance. Le present chapitre tente d'apporter une réponse à une telle question en analysant le risque engendré par le report d'une maintenance préventive. Le risque évalué est relatif aux durées résiduelles de l'équipement à partir de la date optimale de maintenance préventive.

Dans la littérature, des auteurs ont proposé des travaux d'optimisation de stratégie de maintenance avec retard. Dans ce contexte, Nakagawa et Osaki [99] ont introduit un modèle mathématique tenant compte de la disponibilité des pièces de rechange pour l'optimisation de la stratégie de maintenance de type âge. Tilquin et Cléroux [100] ont également proposé

un modèle mathématique pour l'optimisation de la stratégie de maintenance de type bloc [2]. Cette stratégie suggère un remplacement de l'équipement à la panne (remplacement correctif) et après une période de fonctionnement (remplacement préventif). Pour une telle stratégie, le modèle proposé dans [100] [101] [102] introduit une période d'inactivité pendant laquelle aucun remplacement correctif n'est effectué jusqu'à la date du remplacement préventif. L'objectif du modèle de Tilquin et Cléroux [100] est la réduction du nombre de remplacements préventifs prématurés. Un tel modèle engendre un coût total moyen par unité de temps inférieur à celui engendré par le modèle initialement proposé dans [2].

Les approches de retard de maintenance adoptées dans les travaux cités en haut, se basent essentiellement sur une analyse du coût monétaire par unité de temps. Dans le présent chapitre, nous introduisons une autre approche (complémentaire à celle de [99] [100]) basée sur les durées résiduelles de l'équipement à partir de la date où la maintenance préventive est initialement planifiée. Tenant compte du fait que la distribution de Weibull est la distribution largement exploitée en pratique, notre approche sera introduite en supposant sans difficulté de généralisation que l'équipement admet des durées de vie qui suivent une distribution de Weibull à trois paramètres. Les résultats obtenus permettent de montrer qu'à partir d'un âge donné de l'équipement, il devient difficile de prédire la durée de vie résiduelle de l'équipement. De ce fait par la suite, nous considérons que cet âge limite est la date extrême du report de maintenance préventive. Le risque d'un tel report sera basé sur une fonction qui permet après sa date de maintenance préventive.

Dans ce chapitre, la Section suivante analyse la fiabilité d'un équipement dont les durées de vie sont distribuées suivant une loi de Weibull. La Section 3 introduit le problème de report de maintenance préventive pour de tels équipements.

2 Analyse de la fiabilité et mode de défaillance

Nous rappelons que la fonction de fiabilité d'un équipement dont les durées de vie sont représentées par une variable aléatoire positive X distribuée selon une loi de Weibull est :

$$\begin{aligned} R(t) &= P\left(X > t\right), \\ &= \exp\left(-\left(\frac{t-\gamma}{\eta}\right)^{\beta}\right), \end{aligned} \tag{V.1}$$

où β, η et γ sont respectivement les paramètres de forme, d'échelle et d'origine (position). La fiabilité de l'équipement peut être définie en fonction d'autre paramètre t_α solution de

l'Équation (V.2) :

$$\log\left(-\log\left(R(t)\right)\right) = 0. \tag{V.2}$$

Nous remarquons que le paramètre t_α, contrairement aux paramètres (de forme β, d'échelle η et d'origine γ), est un quantile défini par la probabilité $\alpha = 0,6321$ c'est à dire que $R(t_\alpha) = 1 - 0,6321 = 0,3679$ indépendamment de la distribution de Weibull. En utilisant l'expression de $R(t)$ dans l'Équation (V.2), la valeur de t_α est donnée en fonction de η et γ par :

$$t_\alpha = \eta + \gamma. \tag{V.3}$$

Ainsi, la fonction de fiabilité de l'Équation (V.1) peut s'écrire en fonction de β, η et t_α :

$$R(t) = \exp\left(-\left(\frac{t - t_\alpha}{\eta} + 1\right)^\beta\right). \tag{V.4}$$

Cette expression de la fiabilité sera utilisée pour analyser la fiabilité conditionnelle au delà de la date t_α en fonction de la valeur du paramètre de forme β.

2.1 Cas où le paramètre de forme β est entier

En considérant que le paramètre de forme est un entier et en utilisant la formule du binôme de Newton dans l'Équation (V.4), nous réécrivons la fonction de fiabilité :

$$R(t) = \exp\left(-\sum_{k=0}^{\beta} \binom{\beta}{k} \left(\frac{t - t_\alpha}{\eta}\right)^k\right), \tag{V.5}$$

où $\binom{\beta}{k} = \dfrac{\beta!}{(\beta - k)!k!}$. Alors la fonction fiabilité conditionnelle $R_\alpha(t)$ au delà de t_α $(t > t_\alpha)$ est donnée par :

$$\begin{aligned} R_\alpha(t) &= P\left(X > t | X > t_\alpha\right), \\ &= \frac{P\left(X > t; X > t_\alpha\right)}{P(X > t_\alpha)}, \\ &= \frac{P\left(X > t\right)}{P(X > t_\alpha)}. \end{aligned} \tag{V.6}$$

La combinaison des deux Équations (V.5) et (V.6) conduit à l'expression suivante de $R_\alpha(t)$:

$$R_\alpha(t) = \exp\left(-\sum_{k=1}^{\beta} \binom{\beta}{k} \left(\frac{t-t_\alpha}{\eta}\right)^k\right). \tag{V.7}$$

Cette équation est équivalente à :

$$R_\alpha(t) = \exp\left(-\beta\left(\frac{t-t_\alpha}{\eta}\right) - \frac{\beta}{(\beta-2)!2!}\left(\frac{t-t_\alpha}{\eta}\right)^2 - \cdots - \left(\frac{t-t_\alpha}{\eta}\right)^\beta\right). \tag{V.8}$$

La fiabilité conditionnelle s'écrit dans ce cas :

$$R_\alpha(t) = \prod_{k=1}^{\beta} \exp\left(-\left(\frac{t-t_\alpha}{\eta}\right)^k\right)^{\binom{\beta}{k}}. \tag{V.9}$$

Selon l'Équation (V.9), au delà de la date t_α, l'équipement se comporte comme ayant plusieurs modes de défaillance en concurrence. En d'autres termes, l'équipement est tel que son diagramme bloc de fiabilité admet une structure série composée de $N_s = 2^\beta - 1$ composants. Chaque composant est caractérisé par des durées de vie qui suivent une distribution de Weibull de mêmes paramètres d'échelle η et de position t_α. Le $k^{ième}$ groupe de composants du diagramme bloc de fiabilité admet une distribution des durées dont le paramètre de forme est k.

L'Équation (V.9) peut être écrite sous la forme suivante :

$$R_\alpha(t) = \prod_{k=1}^{\beta} \exp\left(-\left(\frac{t-t_\alpha}{\eta/\binom{\beta}{k}^{1/k}}\right)^k\right). \tag{V.10}$$

Cette écriture permet également de mettre en évidence le comportement concurrentiel de β modes de défaillance au lieu de N_s. Le $k^{ième}$ $(k = 1, \cdots, \beta)$ mode de défaillance a une distribution de Weibull de paramétre de forme k, de paramètres d'échelle $\frac{\eta}{\left(\binom{\beta}{k}\right)^{1/k}}$ et d'origine t_α.

Les deux écritures de la fiabilité de l'équipement Équations (V.9-V.10) permettent de montrer qu'au delà de la date t_α, l'équipement se comporte comme ayant une structure mettant en concurrence plusieurs modes de défaillance. Ces équations montrent aussi la difficulté de prédire la probabilité de survie de l'équipement une fois que l'âge de l'équipement dépasse t_α.

2.2 Cas où le paramètre de forme β est réel positif

Lorsque le paramètre de forme de la distribution est un réel positif la définition de l'ensemble de modes de défaillance équivalent est difficile et ne peut se faire que par encadrement selon une borne inférieure et une borne supérieure. Pour celà, supposons que $m = [\beta]$ est la partie entière du paramètre β et $\beta = m + \epsilon$ avec $0 < \epsilon < 1$. La fonction de fiabilité s'écrit alors :

$$\begin{aligned} R(t) &= \exp\left(-\left(\frac{t-\gamma}{\eta}\right)^{\beta}\right), \\ &= \exp\left(-\left(\frac{t-\gamma}{\eta}\right)^{m+\epsilon}\right), \end{aligned} \tag{V.11}$$

comme $t > t_\alpha$ et $\frac{t-t_\alpha}{\eta} > 0$ alors nous obtenons $\frac{t-t_\alpha}{\eta} + 1 > 1$. De plus, sachant $\gamma = t_\alpha - \eta$ nous avons l'inégalité suivante :

$$\begin{aligned} \left(\left(\frac{t-t_\alpha}{\eta}\right)+1\right)^{m} \leq \left(\left(\frac{t-t_\alpha}{\eta}\right)+1\right)^{m+\epsilon} \leq \left(\left(\frac{t-t_\alpha}{\eta}\right)+1\right)^{m+1}, \\ \left(\left(\frac{t-t_\alpha}{\eta}\right)+1\right)^{m} \leq \left(\left(\frac{t-t_\alpha}{\eta}\right)+1\right)^{\beta} \leq \left(\left(\frac{t-t_\alpha}{\eta}\right)+1\right)^{m+1}. \end{aligned} \tag{V.12}$$

À partir de cette inegalité, nous obtenons :

$$R_\alpha(t, m+1) \leq R_\alpha(t, \beta) \leq R_\alpha(t, m), \tag{V.13}$$

où $R_\alpha(t, m+1)$, $R_\alpha(t, \beta)$, $R_\alpha(t, m)$ sont les fonctions de fiabilité conditionnelle au delà de t_α pour les distributions de Weibull de même paramètres d'origine t_α, d'échelle η et de forme respectivement $m+1$, β et m.

Les Équations (V.9-V.10-V.13) donnent les modes de défaillance équivalents de l'équipement au delà de t_α. Ces modes équivalents informent sur le mécanisme et la difficulté de maîtriser le processus de défaillance de l'équipement au delà de l'âge t_α. Ainsi, la date t_α constitue un âge critique pour de tels équipements parce qu'au delà de cette date la prévision de la défaillance devient peu précise. Par conséquent, les possibilités de retarder les actions de maintenance préventive s'arrêtent à cette date t_α.

3 Retard de maintenance préventive

Nous avons évoqué au début de ce chapitre, les hypothèses sur lesquelles reposent la majorité des travaux de recherche ayant proposer des modèles d'optimisation des stratégies de maintenance préventive. Ces modèles stipulent que les actions de maintenance préventive sont exécutées sans retard à la date précise calculée de sorte à maximiser la disponibilité de l'équipement ou encore à minimiser le coût total moyen par unité de temps engendré par les actions de maintenance corrective et préventive. Cependant, des facteurs tels qu'une mauvaise gestion du stock de pièce de rechange, l'indisponibilité des opérateurs de maintenance ou encore le choix de répondre à une demande fortuite d'un produit dans le marché mais aussi des conditions météorologiques très sévères, tous ces facteurs peuvent entraîner un retard quant à l'exécution des actions de maintenance préventive à la date optimale calculée. Dans ces conditions, la décisions à prendre est d'ajourner l'exécution de la maintenance préventive pour une date ultérieure. Le problème qui se pose est jusqu'à quelle date une telle maintenance peut être retardée. Pour répondre à ce problème nous postulons que le retard ne peut excéder une valeur supérieure à t_α. Ce postulat est motivé par les résultats obtenus dans les sections précédentes. En effet, nous avons montré qu'à partir de la date critique t_α, il devient difficile de prédire le comportement en terme de fiabilité de l'équipement. Pour calculer le risque induit par le retard de la maintenance préventive, nous nous baserons sur une fonction dite de survie potentielle et notée $psf(T_0)$. Cette fonction est définie en fonction de la date initiale T_0 de la maintenance préventive et permet de mesurer le risque lié au report d'une telle maintenance. Cette fonction fait un arbitrage entre le pourcentage de temps d'exploitation espéré et la perte de fiabilité due au report de la maintenance :

$$psf(t_\alpha, T_0) = \frac{\frac{t_\alpha - T_0}{\eta}}{R(T_0) - R(t_\alpha)}. \tag{V.14}$$

Si nous posons $\Delta = t_\alpha - T_0$ comme mesure du retard à l'exécution de la maintenance préventive, dans ce cas, la fonction psf devient en fonction de Δ et t_α :

$$psf(t_\alpha, \Delta) = \frac{\frac{\Delta}{\eta}}{R(t_\alpha - \Delta) - R(t_\alpha)}. \tag{V.15}$$

Dans le cas où nous supposons que les durées de vie de l'équipement suivent une distribution de Weibull alors la fonction psf devient :

$$psf(\beta, \eta, \Delta) = \frac{\frac{\Delta}{\eta}}{\exp\left(-\left(1-\frac{\Delta}{\eta}\right)^{\beta}\right) - \exp(-1)}. \tag{V.16}$$

Posons $\zeta = \dfrac{\Delta}{\eta}$, l'Équation (V.16) donne

$$psf(\beta, \zeta) = \frac{\zeta}{\exp\left(-(1-\zeta)^{\beta}\right) - \exp(-1)}. \tag{V.17}$$

Cette équation montre que le risque induit par le report d'une maintenance préventive à la date t_α dépend de deux paramètres à savoir le paramètre de forme β et le ratio $\zeta = \dfrac{\Delta}{\eta}$ avec $0 < \zeta < 1$. Par la suite nous mesurons l'impact des paramètres β et ζ sur la fonction de survie potentielle psf.

3.1 Effet du paramètre β sur la fonction psf

L'effet de β sur la fonction potentielle de survie psf est évalué par la fonction dérivée du risque psf par rapport au paramètre β pour ζ constant. La fonction dérivée du psf par rapport à β est :

$$\frac{\partial}{\partial\beta} psf(\beta, \zeta) = \frac{\zeta \log(1-\zeta) \exp(\beta \log(\zeta))}{\left(\exp\left(-(1-\zeta)^{\beta}\right) - \exp(-1)\right)^2}. \tag{V.18}$$

Nous savons que $\log(1-\zeta) \leq 0$ parce que le ratio ζ est défini de manière que $0 < \zeta < 1$. Cela implique que la fonction dérivée $\dfrac{\partial}{\partial\beta} psf$ est inférieure à 0. Par conséquent, la fonction psf décroit avec le paramètre de forme β ceci indépendamment de la valeur du ratio ζ. Pour différentes valeurs de β, la Figure V.1 montre les variations de la fonction psf en fonction du ratio ζ. Cette figure montre en effet que pour une valeur fixée du ratio ζ, la fonction psf diminue avec l'augmentation de la valeur de β. Le même résultat est obtenu par la Figure V.2 où la fonction psf est évalué par la variation simultanée de β et ζ. L'ensemble des figures montrent que la fonction psf atteint sa valeur maximale pour $\beta = 1$ avec un ratio $\zeta = 0$. Ce qui veut dire que le potentiel de survie est maximum pour $\beta = 1$ et lorsque T_0 est égale à t_α.

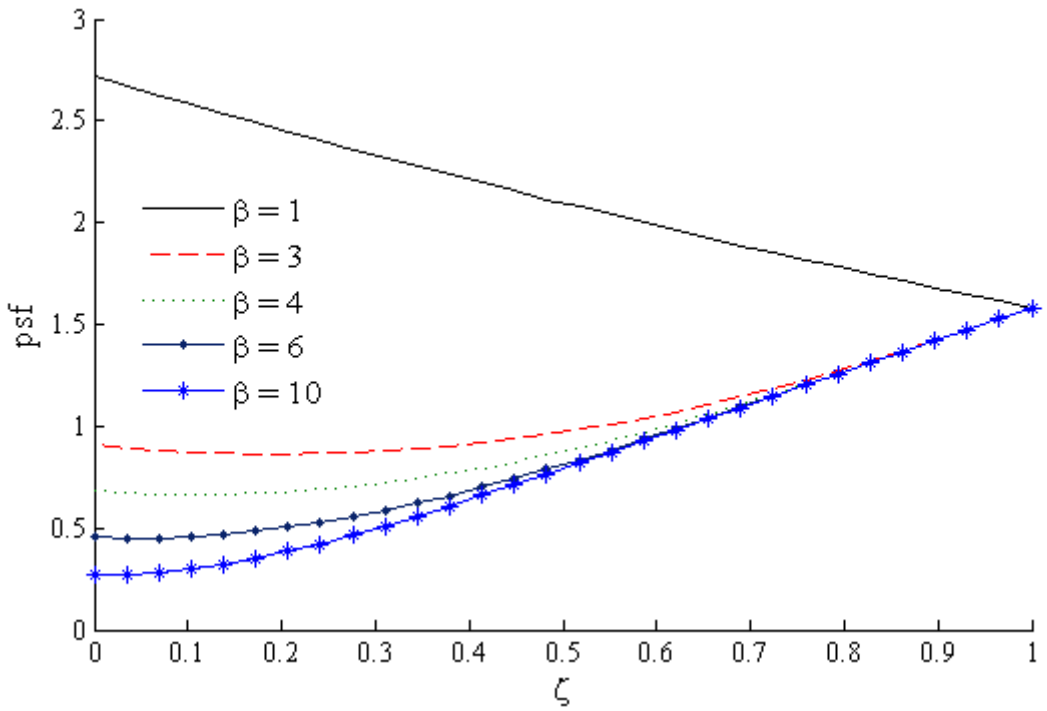

Figure V.1 – Les fonctions de survie potentielle pour les différentes valeurs de β.

3.2 Effet du ratio ζ sur la fonction psf

Par analogie à l'étude précedente, l'effet du ratio ζ sur la fonction risque psf est étudié par l'étude de la dérivée par rapport à ζ de la fonction psf. Pour se faire, nous introduisons une fonction g définie par :

$$g(\zeta) = \exp\left(-(1-\zeta)^{\beta}\right). \tag{V.19}$$

Proposition 11 *La fonction $g(\zeta)$ définie par l'Équation (V.19) est strictement croissante et convexe. De plus pour certaines valeurs de β, il existe $\zeta_0 > 0$ tel que*

$$\frac{dg(\zeta)}{d\zeta} = \frac{g(\zeta) - g(0)}{\zeta}.$$

Preuve. La fonction dérivée de $g(\zeta)$ par rapport ζ est donnée par :

$$\frac{d}{d\zeta}g(\zeta) = \beta\,(1-\zeta)^{\beta-1}\exp\left[-(1-\zeta)^{\beta}\right].$$

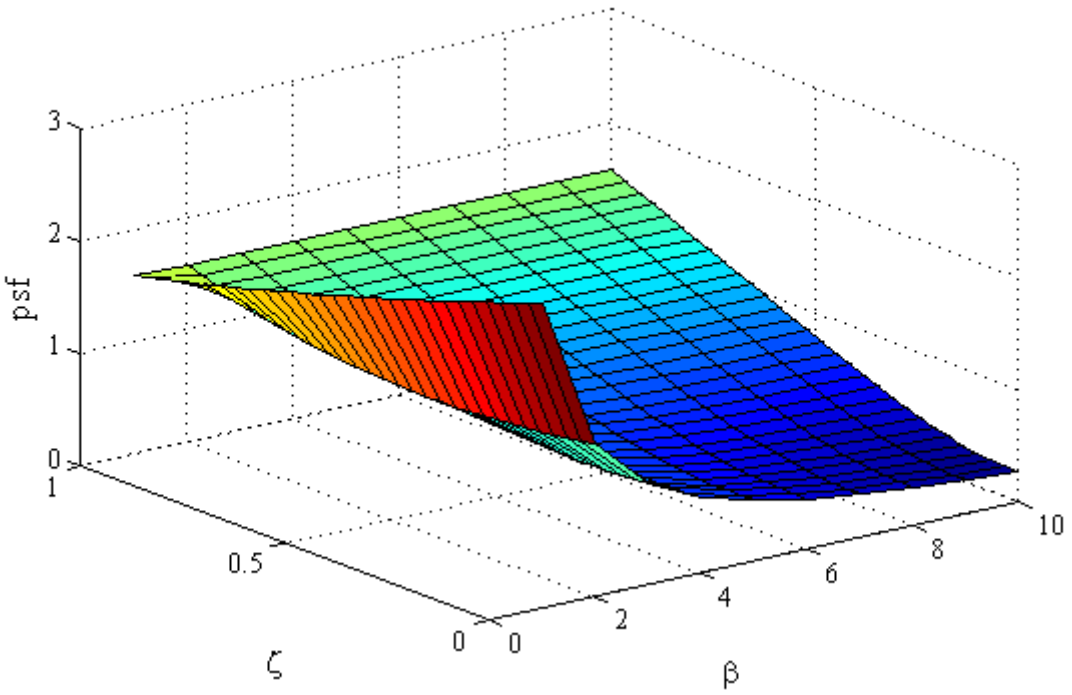

Figure V.2 – Les fonctions de survie potentielle pour les valeurs simultanées de β et de ζ.

Cette fonction dérivée est positive parce que $\zeta \in]0,1[$. Par conséquent, la fonction $g(\zeta)$ est croissante en fonction de ζ. La fonction dérivée seconde de $g(\zeta)$ est :

$$\frac{d^2}{d\zeta^2}g(\zeta) = \beta^2 (1-\zeta)^{\beta-2} \left[(1-\zeta)^{\beta} + 1 - \frac{1}{\beta}\right] \exp\left[-(1-\zeta)^{\beta}\right].$$

De l'expression de la dérivée seconde, nous notons qu'elle est aussi positive pour tout $\beta > 1$ et $0 < \zeta < 1$. Par conséquent, la fonction $g(\zeta)$ est convexe et croissante sur $]0,\zeta]$. Alors pour chaque ζ donnée, il existe un point $\zeta_0 \in]0,\zeta]$ tel que la variation moyenne de $g(\zeta)$ sur $]0,\zeta]$ soit égale à la variation de $g(\zeta)$ au point ζ_0. C'est à dire :

$$\frac{dg(\zeta)}{d\zeta}|_{\zeta=\zeta_0} = \frac{g(\zeta_0) - g(0)}{\zeta_0}.$$

■

En se basant sur la fonction $g(\zeta)$, la fonction psf est donnée par :

$$psf(\zeta) = \frac{\zeta}{g(\zeta) - g(0)}, \tag{V.20}$$

La dérivée première de la fonction psf par rapport à ζ est telle que :

$$\frac{d}{d\zeta}psf(\zeta) = \frac{(g(\zeta) - g(0)) - \zeta\frac{dg(\zeta)}{d\zeta}}{(g(\zeta) - g(0))^2}, \tag{V.21}$$

$$= \frac{1}{g(\zeta) - g(0)}\left(1 - \zeta\frac{\frac{dg(\zeta)}{d\zeta}}{g(\zeta) - g(0)}\right). \tag{V.22}$$

Le signe de la fonction dérivée du psf dépend du terme entre parenthèse à droite de l'égalité dans l'Équation (V.22) car $g(\zeta) - g(0)$ est toujours positif. Il s'ensuit que la fonction $\frac{d}{d\zeta}psf(\zeta) = 0$ admet une solution si

$$\frac{dg(\zeta)}{d\zeta} = \frac{g(\zeta) - g(0)}{\zeta}, \tag{V.23}$$

alors l'Équation (V.23) est identique à

$$g(\zeta_0)\left(1 - \beta\zeta_0(1 - \zeta_0)^{\beta-1}\right) = g(0). \tag{V.24}$$

L'Équation (V.24) a une solution triviale correspondant à $\zeta_0 = 0$. Et d'après les résultats de la Proposition 11, il existe ζ_0, une valeur particulière du ratio ζ qui annule la dérivée $\frac{dpsf(\zeta)}{d\zeta}$. Par conséquent, à $\zeta = \zeta_0$, psf admet un minimum et donc elle est décroissante dans l'intervalle $[0, \zeta_0]$ et elle est croissante dans l'intervalle $[\zeta_0, 1]$. Pour une valeur de $\beta = 2$, le Tableau V.1 donne la valeur de $\zeta_0 = 0,4277$. La Figure V.3 donne les variations de la fonction psf en fonction des ratio ζ et pour $\beta = 2$. Par conséquent, une maintenance préventive peut être retardée si l'on se trouve dans l'intervalle $[0, \zeta_0]$. En effet, lorsque la maintenance préventive est prévu à ζ_i dont l'âge correspondant est $t_i = t_\alpha - \zeta_i\eta$ alors la maintenance préventive peut être retardée jusqu'à une date $t_f = t_\alpha - \zeta_f\eta$ avec $0 < \zeta_f < \zeta_i < \zeta_0$. À titre d'exemple lorsque la distribution de la durée de vie de l'équipement est $W(2, \eta, \gamma)$ alors le premier intervalle s'étale sur $42,77\%\eta$ de l'âge critique. En conséquence, pour un équipement dont la durée de vie suit une loi $W(2, \eta, \gamma)$, il devient intéressant d'attendre l'âge critique pour l'exécution de la maintenance préventive lorsque l'âge de l'équipement dépasse $((57,23\%\eta) + \gamma)$ sans défaillance voir la Figure V.3. La marge $[(57,23\%\eta) + \gamma, t_\alpha]$ devient un temps d'exploitation ou de production supplémentaire. Par conséquent, le report de la maintenance s'intègre logiquement pour réduire les pertes de temps de production dues aux décisions précipitées d'exécution de maintenance préventive. Alors la gestion intégrée du report augmente la durée de production et réduit en même temps le nombre d'actions préventives. Cependant, dans l'intervalle $[\zeta_0, 1]$, la variation de la fonction psf suggère de ne pas reporter la maintenance préventive du fait que le risque de défaillance de l'équipement est de plus en plus important.

Dans le but de permettre l'exploitation de l'approche proposée, la Figure V.4 donne un abaque permettant pour les valeurs données du couple de paramètre β et ζ, d'évaluer la fonction potentielle de survie. La Figure V.4 représente les courbes iso-psf pour les valeurs 0,5 ; 1,0 ; 1,5 ; 2 et 2,5. Chaque courbe représente la même valeur de la fonction psf. Cet abaque est en effet un outil graphique d'aide à la décision quant au report de la maintenance préventive. À titre d'exemple, nous observons que pour des valeurs de $\beta < 1,5$, le report de la maintenance préventive s'accompagne d'une augmentation de la valeur de la fonction de survie potentielle. En effet, il suffit de se reférer à la Figure V.4, de considérer ensuite une ligne verticale passant par le point $\beta = 1,5$ et d'en déduire le résultat.

Tableau V.1 – Les valeurs ζ_0 pour les distributions de Weibull de paramètre de forme β

β	2	3	4	5	6
ζ_0	0.4277	0.1894	0.1045	0.0658	0.0451

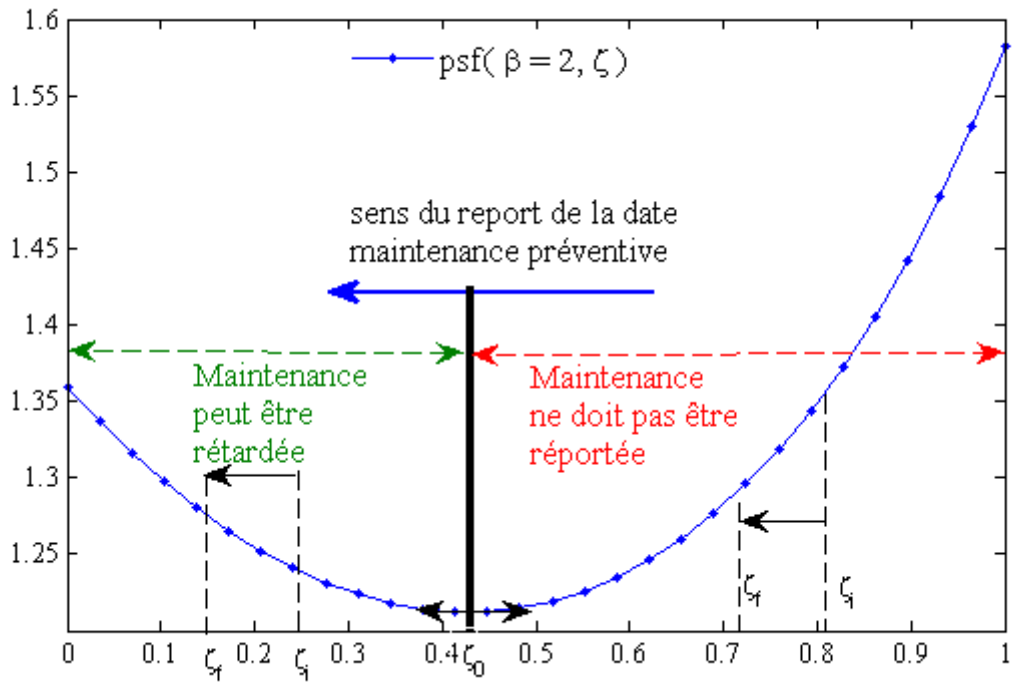

Figure V.3 – Analyse de la fonction de survie potentielle en fonction de ζ pour la distribution $W(2, \eta, \gamma)$.

Les résultats de cette section montrent les limites du report ou du retard de la maintenance préventive Figure V.1. En effet, la variation du psf fait souvent que l'opportinuité que représente le report de la maintenance préventive est faible comparée aux risques de défaillance engendrés.

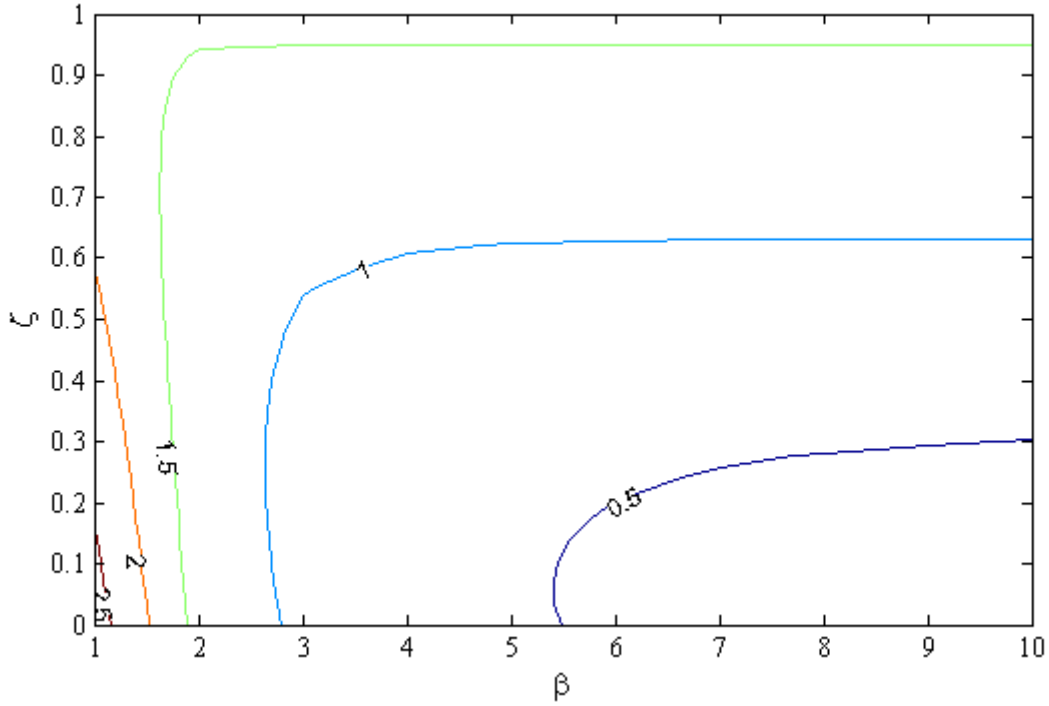

Figure V.4 – Abaque des iso-psf et lignes de niveau de fonction de survie potentielle.

Ce problème apparait dans le cas d'une distribution de Weibull, lorsque le paramètre β est assez élévé ou avec un rapport ζ proche de 1. Cependant, le retard peut être imposé par les contraintes externes comme la demande du marché. Ainsi, nous pouvons réduire les pertes en acceptant un report t_d inférieure à la valeur de t_α. Pour ce cas précis, le report s'arrête à t_d tel que $t_d = t_\alpha - d\eta\zeta$ avec $0 < d < 1$. La fonction psf qui permettra d'évaluer t_d est définie en fonction du triplet (β, ζ, d) :

$$psf(\beta, \zeta, d) = \frac{d\zeta}{\exp\left(-(1-d\zeta)^\beta\right) - \exp\left(-1\right)}. \tag{V.25}$$

La valeur de t_d permet de retarder la date de maintenance préventive sous les conditions défavorables pour faire face aux contraintes de la production sans atteindre l'âge critique de l'équipement.

4 Conclusion

Dans ce chapitre, nous avons proposé une analyse quant aux possibilités de report de la maintenance préventive. Dans cette analyse, nous nous sommes plus intéressés au cas parti-

culier où la distribution des durées de vie est une loi de Weibull. Le risque lié au report de la maintenance est mesuré par une fonction de survie potentielle psf. Cette fonction est complètement définie par le paramètre de forme β de la distribution de Weibull ainsi que le gain en pourcentage de temps d'exploitation ζ. Ce dernier définit le rapport entre la durée supplémentaire d'exploitation et le paramètre d'échelle η de la distribution de Weibull. La durée supplémentaire répresente l'écart entre la date initiale de maintenance préventive et la date critique de fonctionnement de l'équipement.

Dans ce chapitre, nous avons également analysé la fonction psf en fonction des paramètres β et ζ. De cette analyse, nous avons distingué deux intervalles dont les longueurs sont définies en fonction du paramètre de forme de la distribution de Weibull. Sur l'un de ces intervalles, le report de la maintenance préventive demeure possible avec un minimum de risque. Cependant, sur le second intervalle, ce report s'accompagne d'une forte probabilité de défaillance de l'équipement due à la forte variation de la fonction psf avec l'âge de l'équipement. Dans ce cas, la perte de fiabilité devient plus importante que le gain en pourcentage de temps d'exploitation. De ce fait, l'arrêt de l'équipement peut être une solution en attendant l'exécution de la maintenance préventive. Le report de la maintenance préventive n'exclut pas l'occurrence de défaillance. Par conséquent, l'équipement peut en effet tomber en panne pendant le report de maintenance préventive.

L'analyse proposée dans ce chapitre est facilement applicable à d'autres lois de défaillance continues telles que Gamma, Normale tronquée ou un mélange de lois de défaillance. Lorsque les durées de vie d'un équipement sont régies par de telles lois de probabilité, cet équipement peut ne pas avoir un âge critique de fonctionnement. Dans ce cas de figure, la fonction de survie potentielle proposée sera définie par exemple entre deux dates successives de maintenance préventive. La décision de report de la maintenance préventive sera tributaire de la monotonie de la fonction de survie potentielle psf. En effet, le report sera possible sur des intervalles où la perte de fiabilité est moins importante que le pourcentage de gain de temps d'exploitation. Les résultats obtenus et exposés dans ce chapitre ont fait l'objet de publication dans [103].

Conclusion générale et perspectives

L'objectif de ce travail de thèse consiste en le dévelopement de modèles statistiques et d'outils efficaces pour l'analyse des politiques de maintenance basées sur les données de durées de vie hétérogènes. Ce travail de recherche propose en effet une démarche complète d'analyse de stratégies de maintenance en partant des données de durées de vie jusqu'à l'obtention de la politique optimale de maintenance en passant par une phase d'estimation des lois de probabilité. Le rapport de ce travail de thèse est construit autour de cinq chapitres.

Dans le Chapitre I, nous avons posé le problème de l'hétérogénéité dans les données des durées de vie. Dans ce même chapitre, les différentes sources d'hétérogénéité sont identifiées et classées en fonction des modéles de probabilité associés aux données des durées de vie.

Le Chapitre II a permis d'exposer l'ensemble des approches statistiques permettant une estimation des modèles de probabilité basée sur les données des durées de vie. Il s'agit des approches d'estimation paramétrique, semi-paramétrique et non-paramétrique. Dans ce même chapitre, la construction et l'efficacité de chacune des approches sont discutées et comparées. Nous avons montré que la démarche non-paramétrique induit moins de sources d'erreurs et une implémentation facile dans le cadre de données hétérogènes que l'approches paramétrique ou semi-paramétrique. Nous y avons conduit une comparaison entre les méthodes d'estimation de la classe non-paramétrique. À la suite de cette comparaison, la méthode du Noyau de Parzen a été choisie pour l'estimation dans nos travaux de thèse.

Dans le Chapitre III, nous avons décrit la construction et analysé les erreurs de l'estimateur du Noyau de Parzen. Nous avons mis en évidence que seul le paramètre de lissage influe sur la convergence de l'estimation selon les travaux de la littérature. Dans ce même chapitre, l'effet de la variabilité du paramètre de lissage sur l'efficacité des estimations basées sur la méthode du Noyau de Parzen est étudié. Nous avons montré qu'il est possible d'annuler l'effet de la variabilité du paramètre de lissage de la méthode du Noyau sur les résultats d'estimation lorsque la taille des données est conséquente.

Dans le Chapitre IV, nous nous sommes intéressés à l'analyse des politiques de maintenance d'équipements usagés évoluant dans plusieurs environnements. Nous avons évalué et analysé le coût unitaire des stratégies de maintenance de type âge sur les équipements usagés changeant d'environnement d'exploitation. Une analyse qui a permis d'établir les conditions nécessaires et suffisantes garantissant un âge unique de remplacement optimal. Dans ce même chapitre, le coût de cette stratégie de maintenance est estimé par la méthode du Noyau de Parzen. Sur la base des simulations, nous avons montré que les résultats basés sur l'estimation du Noyau de Parzen sont peu variables et proches des valeurs théoriques. Cependant, cette méthode d'estimation nécessite un nombre important de données de durées de vie pour assurer une faible variabilité et convergence des résultats statistiques. La taille démeure une contrainte pour le calcul et la faible variabilité du paramètre de lissage. Cette contrainte pourrait être relaxée avec une connaissance a priori de la distribution statistique du paramètre de lissage. La connaissance a priori de cette distribution conduira à la détermination de sa loi a posteriori et sa valeur la plus probable. Cette nouvelle approche d'estimation nécesssite une combinaison de deux approches statistiques différentes paramétrique et non-paramétrique. Enfin, nous pensons que ces nouveaux modèles permettront de pousser l'introduction de l'analyse non-paramétrique dans le domaine de l'optimisation de la maintenance basée sur les données et garantir des résultats statistiques cohérents avec la réalité.

Par la suite, nous avons établi et étudié des extensions de la stratégie de maintenance proposée. Dans l'une des extensions, nous avons analysé le modèle avec réparation minimale en cas de défaillance dans le premier environnement tandis que l'équipement serait remplacé dans le second environnement. Dans une seconde extension, le modèle proposé tient compte d'une réparation avant le début des opérations dans le second environnement. L'objectif de ces nouvelles analyses était d'établir de manière théorique les conditions nécessaires d'existence d'un âge optimal de remplacement garantissant un coût unitaire minimal.

Le Chapitre V a discuté du retard dans l'exécution de la maintenance de la maintenance préventive. Nous avons proposé une réponse à la question du report de la maintenance préventive. Ce report de la maintenance pouvant être dû à l'indisponibilité de pièces de réchange, d'équipe d'intervention ou des conditions d'exploitation. Ce report fut analysé à travers une fonction de risque. Cette analyse a permis l'identification de deux intervalles. Ainsi, dans l'un des intervalles le report demeure possible avec peu de risque de défaillance pour l'équipement.

Cependant, les politiques de maintenance proposées et analysées, dans ce chapitre, supposent un horizon d'optimisation infini. Une hypothèse qui facile la modélisation mathématique des coûts unitaires de maintenance. En effet, la logique de modélisation sur l'horizon infini émet l'hypothèse que le comportement cyclique moyen du processus de maintenance est le même que

son comportement sur un cycle de fonctionnement. Cette hypothèse induit une interprétation contraignante pour le déroulement opérationnel des cycles de maintenance c'est à dire que les cycles de maintenance doivent être identiques dans leur déroulement. Cette interprétation est implicitement intégré dans les modèles de coûts de nos politiques de maintenance. Pour ces politiques de maintenance, il suffit de considérer un remplacement avec un ensemble d'équipements usagés de même âge T_1 provenant du même environnement.

Comme extensions possibles de ce travail, nous jugeons intéressant d'explorer les points suivants en proposant un cadre statistique permettant d'étudier l'optimisation de la maintenance en considérant :

- l'horizon fini,
- des stratégies de maintenance imparfaite,
- des stratégies d'inspection combinées à des stratégies de maintenance.

Ces modèles peuvent ensuite être étendues pour des équipements complexes multi-composants. Il serait également interéssant de valider l'ensemble des résutats obtenus en exploitant des données issues de milieux industriels.

Annexes

Annexe 1 : Mesures d'erreurs

Cette partie présente les mesures de distance que l'on peut optimiser afin de trouver les meilleurs paramètres pour l'implémentation de la méthode du Noyau. La littérature fournit les mesures de distance définies dans L^q $(q = 1, 2, \cdots)$, la distance Kullback-Leibler et de la vraisemblance ou son logarithme. Ces mesures d'erreur peuvent aussi être regrouper en critères global et local selon qu'elles fournissent un paramètre pour chaque donnée ou constant pour toutes les données de l'échantillon.

1 Critères locaux

Mean Absolute Error MAE dans le L^1

$$MAE(t) = E|\hat{f}_h(t) - f(t)|. \tag{1}$$

Mean Square Error dans L^2

$$MSE(t) = E\left(\hat{f}_h(t) - f(t)\right)^2. \tag{2}$$

2 Critères glocaux

Integral Absolute Error (IAE) : pour les fonctions de densié absolument intégrables (dans L^1)

$$IAE = \int_0^\infty |\hat{f}_h(t) - f(t)|\mathrm{d}t. \tag{3}$$

Mean Integral Absolute Error (MIAE) : pour les fonctions de densié absolument intégrables (dans L^1)

$$MIAE = E\left(\int_0^\infty |\hat{f}_h(t) - f(t)|\mathrm{d}t\right). \tag{4}$$

Integral Square Error ISE dans L^2

$$ISE = \int_0^\infty \left(\hat{f}_h(t) - f(t)\right)^2 \mathrm{d}t. \tag{5}$$

Annexe 1 : Mesures d'erreurs

Mean Integral Square Error (MISE) : pour les fonctions de densié de carré intégrable (dans L^2)

$$MISE = E\left(\int_0^\infty \left(\hat{f}_h(t) - f(t)\right)^2 \mathrm{d}t\right). \tag{6}$$

Mesure de Kullback-Leibler

$$KL\left(f, \hat{f}_h\right) = \int_0^\infty f(t) \log\left(\frac{f(t)}{\hat{f}_h(t)}\right) \mathrm{d}t. \tag{7}$$

Mesure de Vraisemblance (MV)

$$\begin{aligned} MV\left(\hat{f}_h\right) &= \prod_{i=1}^n \hat{f}_h(t_i), \\ \log\left(MV\right) &= \sum_{i=1}^n \log\left(\hat{f}_h(t_i)\right). \end{aligned} \tag{8}$$

Annexe 2 : Autres méthodes de calcul du paramètre de lissage h

3 Vraisemblance

La technique de détermination par vraisemblance est une imitation de la vraisemblance paramétrique. La mesure de la vraisemblance ($MV(h)$) est obtenue par le produit de l'estimation de la densité à chaque t_i de l'échantillon par la méthode du Noyau de Parzen :

$$\begin{aligned} MV(h) &= \prod_{i=1}^{n} \hat{f}_h\left(t_i\right), \\ &= \prod_{i=1}^{n} \frac{1}{nh} \sum_{j=1}^{n} K\left(\frac{t_j - t_i}{h}\right). \end{aligned} \tag{9}$$

Cette formulation bien que simple de par sa construction [65], ne semble pas être efficace en pratique. Cette inefficacité est intuitive, car imaginer un échantillon de données de durées de vie $i.i.d$ avec des données aberrantes (où la densité de panne devrait être nulle). La technique de vraisemblance exclut pratiquement les paramètres de lissage qui fournissent une fenêtre de lissage annulant la valeur de l'estimation de la densité en ces observations aberrantes de l'échantillon. Car pour ces paramètres de lissage la fonction de vraisemblance serait nulle. Par conséquent, elle aura tendance à proposer une fenêtre de lissage toujours plus large que l'optimale (i.e. que celle attendu) et conduit à une perte d'information avec un biais fort et une courbe de densité de panne beaucoup plus lisse.

L'efficacité de cette technique en estimation paramétrique résulte de la définition a priori de la loi sous-jacente que l'on suppose être la loi source des données [104]. Une loi de distribution qui prend en compte la structure de la probabilité d'apparition des données. Par contre, dans l'estimation non-paramétrique l'absence d'expression de la fonction de densité pour les données induit une sous-information statistique dans les calculs de la vraisemblance. Par conséquent, la méthode basée sur la seule information des données n'est pas protégée contre les données aberrantes. Même s'il est possible de réduire cet effet par validation croisée.

Duin [65] montra que l'utilisation de l'estimateur obtenu par validation croisée apporte de meilleures améliorations à la méthode du maximum de vraisemblance surtout pour les cas

d'une distribution multimodale. Dans ce cas de figure, la fonction de vraisemblance devient après application de la validation croisée :

$$MV(h) = \prod_{i=1}^{n} \frac{1}{nh} \sum_{j=1, i \neq j}^{n} K\left(\frac{t_i - t_j}{h}\right). \tag{10}$$

4 Minimisation dans L^1 (densité absolument intégrable)

L'étude des erreurs dans L^1 a été en grande partie réalisée par Devroye. Les erreurs dans L^1 sont beaucoup proches de la réalité, car elle mesure un écart direct ou sa moyenne. Cependant sa mise en place pratique est difficile mathématiquement, pour l'estimation. Cette difficulté est due à la présence de la valeur absolue, très complexe à manipuler mathématiquement [55], [105].

Sous les mêmes conditions que dans L^2, Devroye a montré la convergence en probabilité de l'estimateur de Parzen dans L^1 et de sa moyenne. Par la suite Devroye proposa une borne inférieure pour la moyenne de l'erreur absolue [57]. Il a montré dans [106] l'indifférence du choix du noyau sur le paramètre de lissage et produit deux résultats essentiels. Le premier est la convergence vers zéro de l'erreur moyenne globale avec une dispersion faible (presque surement) en fonction de l'échantillon. Le second est la C-optimalité de l'estimateur dans L^1, surtout pour les noyaux spécifiques à voir dans [106].

La convergence des mesures dans L^1 est faible car exprimée en probabilité et en plus cette analyse ne fournit pas une méthodologie de calcul du paramètre de lissage. La démarche dans L^1 demeure une analyse plutôt théorique [63] sur la convergence de la méthode du Noyau de Parzen. Car il engendre des calculs complexes.

Dans [105], les auteurs ont montré une équivalence presque entre l'optimisation des erreurs dans L^1 et L^2 de manière locale et globale. On déduit de leur étude que le paramètre de lissage issu de L^1 converge plus rapidement vers zéro que celui L^2. Par contre $n * h$ (taille*paramètre de lissage) le produit par la taille du paramètre de lissage issu de L^2 tend plus rapidement vers l'infinie que celui de L^1. Hall et Wand [105] ont proposé un algorithme (plug-in) de calcul du paramètre de lissage local minimisant une approximation de l'espérance de l'écart absolue et ont donné pour quelques distributions connues une valeur asymptotiquement optimale du paramètre de lissage.

Bien qu'efficace cette technique de calcul est limitée car inutilisable en pratique lorsque l'on ignore la fonction de distribution à estimer ce qui est le cas en réalité.

Tout comme la méthode de vraisemblance la minimisation dans L^1 présente des difficultés d'estimation de queues de distribution. Ceci a été montré par Beirland et Devroye [107] en se basant sur les travaux [57] [106]. Ils trouvent une borne inférieure non-nulle entre l'estimation de la densité et la fonction de densité.

Le principal problème avec L^1 est son aspect peu pratique bien que la méthode plug-in présentée dans [105] soit efficace. Elle contient la fonction de densité réelle en plus de sa dérivée seconde. Ces deux fonctions sont estimées à chaque étape pour assurer la convergence vers la fenêtre de lissage optimale. A travers l'analyse comparée des résultats dans L^1, L^2 [63], nous déduisons que les estimations sur la base du paramètre de lissage de L^1 a une variance

importante que celui de L^2. Cependant, son biais est inférieur à celui de L^2. Ces assertions là sont expliquées par l'expression du paramètre de lissage et du produit $n * h$.

5 Rule-and-thumb (règle et pouce)

Parmi toutes les approches de calcul du paramètre, elle demeure la plus simple à mettre en œuvre. Contrairement aux autres méthodes, elle nécessite une hypothèse supplément c'est à dire un a priori sur la loi de distribution comme dans les modèles paramétriques. En supposant que la densité à estimer est une distribution uni-modale proche de celle d'une distribution Normale de même variance $N(m, \sigma)$ [52].

$$\hat{R}g(f") = \int_{-\infty}^{\infty} \left(\frac{1}{\sigma^5 \sqrt{2\pi}} \left(t^2 - \frac{1}{\sigma^2} \exp\left(-\frac{t^2}{2\sigma^2} \right) \right) \right), \tag{11}$$

$$h_{rt} = \left(\frac{8\sqrt{\pi} Rg\,(K)}{3\sigma_K^4} \right)^{1/5} \sigma n^{-1/5}. \tag{12}$$

Le terme entre parenthèse dans l'Équation (12) est une constante par rapport au noyau utilisé ainsi il est égale à $(1,06; 2,34; 2,78)$ respectivement pour les noyaux (Normal, Epanechnikov, Biweight). La variance σ^2 est obtenue en utilisant un estimateur de dispersion. Tels que s^2 sans biais, avec biais ou une statistique basée sur les rangs par exemple la distance inter-quartile IRQ.

- l'estimateur sans biais de la variance $s^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left(t_i - \bar{t} \right)^2$,
- l'estimateur avec biais de la variance $\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} \left(t_i - \bar{t} \right)^2$.

Pour plus de détail sur cette approche, le lecteur intéressé peut se reporter au travail de Deheuvels [52]. Autrement, la distance inter-quartile $IRQ = Q_3 - Q_1$ est aussi utilisée pour des raisons de robustesse relativement aux données extrêmes ou aberrantes. Un choix qui est encouragé par Silverman [44] et Deheuvels [52]. Il est souvent préférable de prendre le plus petit des deux coefficients pour le paramètre de lissage, c'est à dire :

$$h_{rt} = 1,06 Min\{s, \frac{IRQ}{1,39}\} \tag{13}$$

La pertinence de cette technique [44] se limite aux cas des distributions uni-modales et sans une forte asymétrie. Son usage dans le cas bimodal nécessite une correction sur le coefficient. Silverman conseille une réduction de la constante de $1,06$ à $0,9$ dans l'expression de h_{rt} (Equation (13)) lorsque l'on utilise un noyau gaussien.

Ceci rejoint en partie l'analyse de Deheuvels dans [52] sur le choix d'un paramètre de lissage plus petit que la valeur l'optimale garantissant ainsi un faible biais.

6 Mesure basée sur la fonction de répartition

Nous analysons une autre démarche distincte de celles décrites ci-dessus dans laquelle nous nous focaliserons sur une distance relative à la fonction de répartition empirique. Ce paramètre de lissage résulte des travaux de Sarda Alman et Leger [108] [109] et se base sur les statistiques d'ordre associé à l'échantillon des données de durées de vie. Dans cette section, on discutera le cas de paramètres de lissage constants et variables.

6.1 Paramètre de lissage constant

L'erreur mesure la distance par rapport à l'estimateur empirique de la fonction de répartitions : estimée par la méthode du Noyau de Parzen et une estimation empirique basée sur les statistiques d'ordre. Les conditions de convergence sont explicitées dans [108]. Alman et Leger [109] comparèrent les cas avec/sans validation croisée ($i = j$) et prouvèrent que pour cette mesure la différence est infime. Le critère à optimiser en prenant une pondération linéaire est :

$$CE(h) = \sum_{i=1}^{n} \left(\hat{F}_h(t) - F(t)\right)^2, \tag{14}$$

$\hat{F}_h$ est un estimateur non-paramétrique de la fonction de répartition et définit par,

$$\hat{F}_h(t) = \frac{1}{n} \sum_{i=1}^{n} \int_0^t K\left(\frac{u - t_i}{h}\right), \tag{15}$$

et $\hat{F}_n$ est un estimateur empirique basé sur les statistiques d'ordre de la fonction de répartition calculée à partir des rangs moyen, médian, CDF [1] ou aléatoire. Ainsi la fonction de répartition empirique est définie par :

$$\hat{F}_n(t) = \frac{i - \epsilon}{n + 1 - 2\epsilon} \text{ avec } t_{(i)} < t < t_{(i+1)}, \tag{16}$$

ϵ détermine le type d'estimateur de la fonction de répartition. C'est-à-dire $\epsilon = 0$ rang moyen, $\epsilon = 0.3$ médian, $\epsilon = 0.5$ CDF.

Cette fonction de répartition empirique est définie en escalier et est continue par morceau. Cet estimateur empirique de la fonction de répartition a une distribution Normale [14] et une variance égale à $\frac{F(t)(1-F(t))}{n}$ à constante près. Cette variance tend vers zéro avec la taille des données (alors $\hat{F}_n$ est consistante). Par conséquent, $\hat{F}_n$ tend vers F la fonction de répartition réelle lorsque n devient grand alors par transitivité l'estimation $\hat{F}_h$ basée sur la fonction noyau par minimisation $CE(h)$ tend aussi vers F.

En prenant en compte la continuité de la fonction noyau nous allons modifier la fonction d'erreur proposée par Sarda de manière à garantir la continuité sur chaque intervalle $[t_{(i)}; t_{(i+1)}[$. La modification permet de minimiser la différence d'aire sous les courbes estimées par la méthode de Parzen et les escaliers obtenus de manière empirique sur chaque sous-intervalle. Le

1. cumulative distribution function

critère devient après modification.

$$EC(h) = \sum_{i=1}^{n-1} \int_{t_{(i)}}^{t_{(i+1)}} \left(\hat{F}_h(t) - \hat{F}_{i,n}(t)\right)^2 \mathrm{d}t, \tag{17}$$

où $\hat{F}_{i,n}(t) = p\hat{F}_n(t_{(i)}) + (1-p)\hat{F}_n(t_{(i+1)})$ si $t_{(i)} < t < t_{(i+1)}$ avec $0 \leq p < 1$. Ainsi le paramètre de lissage optimal permet d'annuler la derivée première de la fonction $EC(h)$.

$$\frac{\partial}{\partial h} EC(h) = 0,$$

$$\sum_{i=1}^{n-1} \int_{t_{(i)}}^{t_{(i+1)}} \left(\frac{1}{h}\hat{F}_h(t) - \frac{1}{n}\sum_{i=1}^{n} \frac{t-t_i}{h} K\left(\frac{t-t_i}{h}\right)\right) \left(\hat{F}_h(t) - \hat{F}_{i,n}(t)\right) \mathrm{d}t = 0. \tag{18}$$

En utilisant l'intégration numérique de Simpson dans l'Équation (18), on arrive à atteindre la valeur optimale du paramètre lissage pour cette mesure par la méthode des sécantes [21] évitant le calcul de la dérivée seconde de $EC(h)$ dans la methode de Newton-Raphson.

6.2 Paramètre de lissage variable

Tout au long de ce rapport nous avons expliqué les techniques permettant le calcul des paramètres de lissage constants pour l'ensemble des données. D'autres approches existent et conduisent à des paramètres de lissage variables. On distingue deux techniques proposant des paramètres de lissage variables : en fonction des données de référence (c'est à dire que pour chaque t_i on détermine un paramètre de lissage $h = h(t_i)$) et en fonction du point d'estimation (c'est à dire $h = h(t)$ où l'on veut estimer la densité).

6.2.1 Paramètre $h(t_i)$ dépend des t_i

La méthode que nous proposons est une extension de la méthode basée sur l'écart des fonctions de la répartition proposée ci-dessus. Dans ce cas on cherche un vecteur de paramètres de lissage de mêmes tailles que les données et à chaque donnée on associe un paramètre de lissage $H = (h_1, h_2, \cdots, h_n)$ comme dans le SVM.

$$EC(H) = \sum_{m=1}^{n-1} \int_{t_m}^{t_{m+1}} \left(\hat{F}_H(t) - \hat{F}_n(t)\right)^2 \mathrm{d}t. \tag{19}$$

Le vecteur des paramètres de lissage est obtenu en utilisant une méthode de descente. L'avantage de la recherche d'un paramètre de lissage local est bien vu mais nécessite un temps de calcul énorme pour peu d'efficacité au niveau de l'amélioration des critères. Elle devient inadaptée lorsque la taille de l'échantillon augmente .

6.2.2 Paramètre $h(t)$ en fonction de t

Une approche similaire est le critère MSE. Une approche similaire a été étudiée par Hall et Schucany dans [110]. Les auteurs réajustent la technique de validation croisée UCV par rapport au voisinage du point d'estimation t. L'application d'une telle technique fait face à des difficultés de résolutions. Ainsi dans ce cas précis avec les fonctions de répartition, on aura besoin d'une fonction noyau invariante par translation afin de faciliter les calculs. Cette approche ne sera pas étudiée mais elle est juste montrée ici comme une possibilité. L'expression de la fonction répartition correspondant à cette approche s'écrit :

$$\hat{F}_{h(t)}(t) = \frac{1}{n}\sum_{i=1}^{n}\int_0^t \frac{1}{h(u)} K\left(\frac{u-t_i}{h(u)}\right) \mathrm{d}u. \tag{20}$$

Références bibliographiques

[1] M. Zelen. *Statistical theory of reliability : proceedings of an advanced seminar*. University of Wisconsin Press (1963). 1, 10, 80

[2] R.E. Barlow and F. Proschan. *Mathematical Theory of Reliability*. Society for Industrial and Applied Mathematics (1987). 1, 80, 113

[3] H. Kaffel. Thèse de Doctorat, Université de Laval, Laval Canada (2001). 7

[4] A. ALHOUAIJ. *Contribution à l'optimisation de la maintenance dans un contexte distribué*. Thèse de Doctorat, Grenoble INPG, Grenoble France (2010). 7

[5] P. Ribot. *Vers l'intégration diagnostic-pronostic pour la maintenance des systèmes complexes*. Thèse de Doctorat, Université de Toulouse, Toulouse France (2009). 7

[6] J.F. Lawless. *Statistical models and methods for lifetime data*. Wiley series in probability and mathematical statistics. Wiley (1982). 8, 9, 10, 20, 24

[7] H. Bertholon. *Une modélisation du vieillissement*. Thèse de Doctorat, Université de Joseph Fourier, Grenoble France (2001). 10

[8] W.A. Thompson. *On the foundations of reliability*. Technometrics **23**(1), 1–13 (1981). 14

[9] N.R. Mann, R. E. Schafer, and N.D. Singpurwalla. *Methods for statistical analysis of reliability and life data*. Wiley (1974). 16

[10] H. Procaccia, E. Ferton, and M. Procaccia. *Fiabilite et maintenance des materiels industriels reparables et non reparables*. Tec et Doc (2011). 16

[11] A. Lannoy and H. Procaccia. *Livre Evaluation et maîtrise du vieillissement industriel*. Lavoisier (2005). 16

[12] A. Lannoy and H. Procaccia. *Evaluation de la fiabilité prévisionnelle.* Lavoisier (2006). 16

[13] J.F. Lawless. *Statistical methods in reliability.* Technometrics **25**(4), 305–316 (1983). 16, 80

[14] G. Saporta. *Probabilités : analyse des données et statistique.* Technip (1990). 20, 27, 36, 136

[15] J.I. McCool. *Inference on weibull percentiles and shape parameter from maximum likelihood estimates.* IEEE Transactions on Reliability **R-19**(1), 2–9 (1970). 23

[16] D. Markovic, D. Jukic, and M. Bensic. *Nonlinear weighted least squares estimation of a three-parameter weibull density with a nonparametric start.* Journal of Computational and Applied Mathematics **228**(1), 304–312 (2009). 24, 25

[17] D. Cousineau. *Fitting the three-parameter weibull distribution : review and evaluation of existing and new methods.* IEEE Transactions on Dielectrics and Electrical Insulation **16**(1), 281–288 (2009). 24, 25

[18] B. Abbasi, Abdol H. Eshragh J., J. Arkat, and M. Hosseinkouchack. *Estimating the parameters of weibull distribution using simulated annealing algorithm.* Applied Mathematics and Computation **183**(1), 85–93 (2006). 24, 25

[19] H. Qiao and C.P. Tsokos. *Parameter estimation of the weibull probability distribution.* Mathematics and Computers in Simulation **37**(1), 47–55 (1994). 24, 25

[20] H. Qiao and C.P. Tsokos. *Estimation of the three parameter weibull probability distribution.* Mathematics and Computers in Simulation **39**(1-2), 173–185 (1995). 24, 25

[21] A. Fortin. *Analyse numéique pour ingénieurs.* Presses internationales Polytechnique (2008). 24, 56, 137

[22] D. Markovic and D. Jukic. *On nonlinear weighted total least squares parameter estimation problem for the three-parameter weibull density.* Applied Mathematical Modelling **34**(7), 1839–1848 (2010). 24, 25

[23] D. Jukic and D. Markovic. *On nonlinear weighted errors-in-variables parameter estimation problem in the three-parameter weibull model.* Applied Mathematics and Computation **215**(10), 3599–3609 (2010). 24, 25

[24] T. Bucar, M. Nagode, and M. Fajdiga. *Reliability approximation using finite weibull mixture distributions.* Reliability Engineering & System Safety **84**(3), 241–251 (2004). 24

[25] B. Veber, M. Nagode, and M. Fajdiga. *Generalized renewal process for repairable systems based on finite weibull mixture.* Reliability Engineering & System Safety **93**(10), 1461–1472 (2008). 24

[26] G.J. McLachlan and T. Krishnan. *The EM algorithm and extensions.* Wiley (1997). 24

[27] D. Foata and A Fuchs. *Calcul des probabilités : cours et exercices corrigés.* Masson, Paris (1996). 25

[28] M. Krifa. *A mixed weibull model for size reduction of particulate and fibrous materials.* Powder Technology **194**(3), 233–238 (2009). 26

[29] B. Kleine and B. Bertsche. *Estimating parameters of mixed weibull distributions using genetic algorithms.* Reliability Risk and safety Ale Papazoglou&Zio(eds) pages 861–867 (2010). 26

[30] J.J. Droesbeke, B. Fichet, and P. Tassi. *Analyse statistique des durées de vie : modélisation des données censurées.* Économica. (1989). 26, 27

[31] D.R. Cox. *Regression models and life-tables.* Journal of the Royal Statistical Society. Series B (Methodological) **34**(2), 187–220 (1972). 27

[32] A. Zuashkiani, D. Banjevic, and A.K.S Jardine. *Estimating parameters of proportional hazards model based on expert knowledge and statistical data.* Journal of the Operational Research Society **60**(12), 1621–1636 (2008). 27

[33] F. Brissaud, B. Lanternier, D. Charpentier, and P. Lyonnet. In *Actes de la 3ème édition du colloque international francophone Performances et Nouvelles Technologies en Maintenance, PENTOM 2007* (2007). 27

[34] A.K.S Jardine and A.H.C Tsang. *Maintenance, replacement, and reliability : theory and applications.* CRC/Taylor & Francis (2006). 27, 80

[35] Y.P. Mack. *Rate of strong uniform convergence of k-NN density estimates.* Journal of Statistical Planning and Inference **8**(2), 185–192 (1983). 28

[36] D. Bosq and J.P. Lecoutre. *Theorie de l'estimation fonctionnelle.* Economica (1987). 28, 31, 39, 43, 44, 51, 52

[37] Y.P. Mack and M. Rosenblatt. *Multivariate k-nearest neighbor density estimates.* Journal of Multivariate Analysis **9**(1), 1–15 (1979). 28, 29

[38] P. Hall. *On near neighbour estimates of a multivariate density.* Journal of Multivariate Analysis **13**(1), 24–39 (1983). 29

[39] G.S. Watson. *Density estimation by orthogonal series.* The Annals of Mathematical Statistics **40**(4), 1496–1498 (1969). 30

[40] A.A. Farag and R.M. Mohamed. classification of multispectral data using support vector machines approach for density estimation. (2003). 30, 31

[41] R.M. Mohamed and A.A. Farag. Mean field theory for density estimation using support vector. pages 495–501 (2004). 30, 31

[42] P. Hominal and P. Deheuvels. *Estimation non paramétrique de la densité compte-tenu d'informations sur le support.* Revue de statistique appliquee **29**(3), 47–68 (1979). 32, 47, 69

[43] S. Saoudi, F. Ghorbel, and A. Hillion. *Some statistical properties of the kernel diffeomorphism estimator.* Applied Stochastic Models and Data Analysis **13**(1), 39–58 (1997). 32, 48, 51, 75, 91

[44] B.W. Silverman. *Density Estimation for Statistics and Data Analysis.* Chapman and Hall/CRC, 1 edition (1986). 32, 42, 43, 47, 50, 51, 52, 69, 135

[45] M.C. Jones. *Simple boundary correction for kernel density estimation.* Statistics and Computing **3**(3), 135–146 (1993). 32, 47

[46] J.S. Marron and D. Ruppert. *Transformations to reduce boundary bias in kernel density estimation.* Journal of the Royal Statistical Society. Series B (Methodological) **56**(4), 653–671 (1994). 32, 48, 50, 51, 75

[47] A. Cowling and P. Hall. *On pseudodata methods for removing boundary effects in kernel density estimation.* Journal of the Royal Statistical Society. Series B (Methodological) **58**(3), 551–563 (1996). 32, 47, 48

[48] E. Parzen. *On estimation of a probability density function and mode.* The Annals of Mathematical Statistics **33**(3), 1065–1076 (1962). 38, 51

[49] E. Parzen. *Nonparametric statistical data modeling.* Journal of the American Statistical Association **74**(365), 105–121 (1979). 38, 51

[50] M. Samiuddin and G.M. El-sayyad. *On nonparametric kernel density estimates.* Biometrika **77**(4), 865–874 (1990). 42

[51] D.B.H. Cline. *Admissibile kernel estimators of a multivariate density.* The Annals of Statistics **16**(4), 1421–1427 (1988). 42

[52] P. Deheuvels. *Estimation non paramétrique de la densité par histogrammes généralisés.* **25**(2), 5–42 (1977). 42, 135

[53] S. Adjabi and K. Lagha. évaluation de performances des systèmes d'attente fiables, (2010). 42

[54] S. Adjabi and M. Cherfaoui. Bootstrap dans l'estimation de la densité par la méthode du noyau, (2010). 42

[55] L. Devroye. *The equivalence of weak, strong and complete convergence in l1 for kernel density.* **11**(3), 896–904 (1983). 43, 51, 134

[56] L. Devroye. *On the non-consistency of the l2-cross-validated kernel density estimate.* Statistics & Probability Letters **8**(5), 425–433 (1989). 43, 55

[57] L. Devroye. *A universal lower bound for the kernel estimate.* Statistics & Probability Letters **8**(5), 419–423 (1989). 43, 134

[58] M.C. Jones and P.J. Foster. *A simple nonnegative boundary correction method for kernel density estimation.* Statistica Sinica **6**, 1005–1013 (1996). 47, 48

[59] S.X. Chen. *Probability density function estimation using gamma kernels.* Annals of the Institute of Statistical Mathematics **52**(3), 471–480 (2000). 48, 49, 50

[60] S.X. Chen. *Beta kernel estimators for density functions.* Computational Statistics & Data Analysis **31**(2), 131–145 (1999). 48, 49, 50

[61] S. Adjabi, K. Lagha, and M. Hassani. Méthode du noyau dans l'évaluation de performances des systèmes d'attente, (2009). 48

[62] T. Bouezmarni and J.M. Rolin. *Consistency of the beta kernel density function estimator.* Canadian Journal of Statistics **31**(1), 89–98 (2003). 49

[63] M.P. Wand and L. Devroye. *How easy is a given density to estimate.* Computational Statistics & Data Analysis **16**(3), 311–323 (1993). 50, 51, 134

[64] R.J. Karunamuni and T. Alberts. *On boundary correction in kernel density estimation.* Statistical Methodology **2**(3), 191–212 (2005). 50, 51

[65] R.P.W. Duin. *On the choice of smoothing parameters for parzen estimators of probability density functions.* Computers, IEEE Transactions on **C-25**(11), 1175 –1179 (1976). 51, 133

[66] S. Kullback and R.A. Leibler. *On information and sufficiency.* The Annals of Mathematical Statistics **22**(1), 76–86 (1951). 51

[67] P. Hall and M.P. Wand. *Minimizing l1 distance in nonparametric density estimation.* Journal of Multivariate Analysis **26**(1), 59–88 (1988). 51

[68] C.J. Stones. *An asymptotically optimal window selection rule for kernel density estimates.* **12**(4), 1285—1297 (1984). 54

[69] A.W. Bowman. *An alternative method of cross-validation for the smoothing of density estimates.* Biometrika **71**(2), 353–360 (1984). 54

[70] J.S. Marron. *Comments on a data based bandwidth selector.* Computational Statistics & Data Analysis **8**(2), 155–170 (1989). 55, 58

[71] P. Hall and J.S. Marron. *On the amount of noise inherent in bandwidth selection for a kernel density estimator.* The Annals of Statistics **15**(1), 163–181 (1987). 55

[72] W. Feluch and J. Koronacki. *A note on modified cross-validation in density estimation.* Computational Statistics & Data Analysis **13**(2), 143–151 (1992). 55, 56

[73] D.W. Scott and G.R. Terrell. *Biased and unbiased cross-validation in density estimation.* American Statistical Association **82**(400), 1131–1146 (1987). 55, 57

[74] S.J. Sheather. *Density estimation.* Statistical Science **19**(4), 588–597 (2004). 55, 57, 58

[75] B.U. Park and J.S. Marron. *Comparison of data-driven bandwidth selectors.* Journal of the American Statistical Association **85**(409), 66–72 (1990). 55, 58, 59, 60

[76] P. Hall and J.S. Marron. *Estimation of integrated squared density derivatives.* Statistics & Probability Letters **6**(2), 109–115 (1987). 57, 59

[77] SJ Sheather and MC Jones. *A reliable data-based bandwidth selection method for kernel density estimation.* Journal of the Royal Statistical Society. Series B (Methodological) **53**(3), 683–690 (1991). 59, 60, 91

[78] M.C. Jones and S.J. Sheather. *Using non-stochastic terms to advantage in kernel-based estimation of integrated squared density derivatives.* Statistics & Probability Letters **11**(6), 511–514 (1991). 59, 91

[79] R.E. Barlow and L. Hunter. *Optimum preventive maintenance policies.* Operations Research **8**(1), 90–100 (1960). 80

[80] D.I. Cho and M. Parlar. *A survey of maintenance models for multi-unit systems.* European Journal of Operational Research **51**(1), 1–23 (1991). 80

[81] R. Dekker. *Applications of maintenance optimization models : a review and analysis.* Reliability engineering & systems safety **51**(3), 229–240. 80

[82] T. Nakagawa and S. Mizutani. *A summary of maintenance policies for a finite interval.* Reliability Engineering & System Safety **94**(1), 89–96 (2009). 80

[83] T. Nakagawa. *Advanced reliability models and maintenance policies.* Springer (2008). 80

[84] F.P.A. Coolen, P. Coolen-Schrijner, and K.J. Yan. *Nonparametric predictive inference in reliability.* Reliability Engineering & System Safety **78**(2), 185–193 (2002). 81

[85] P. Coolen-schrijner and F.P.A. Coolen. *Non-parametric predictive inference for age replacement with a renewal argument. quality and reliablity engineering.* International pages 203–215 (2004). 81

[86] B. Bergman and B. Klefsjo. *A graphical method applicable to age-replacement problems.* IEEE Transactions on Reliability **R-31**(5), 478–481 (1982). 81

[87] B. Bergman and B. Klefsjo. *The TTT-concept and replacements to extend system life.* European journal of operational research **28**(3), 302–307 (1987). 81

[88] R.E. Barlow and R.A. Campo. *Total Time on Test Processes and Applications to Failure Data Analysis.* Defense Technical Information Center (1975). 81

[89] T. Dohi, N. Matsushima, N. Kaio, and S. Osaki. *Nonparametric repair-limit replacement policies with imperfect repair.* European journal of operational research **96**(2), 260–273 (1987). 81

[90] T. Dohi, N. Kaio, and S. Osaki. *A new graphical method to estimate the optimal repair-time limit with incomplete repair and discounting.* Computers & Mathematics with Applications **46**(7), 999–1007 (2003). 81

[91] T. Dohi, A. Ashioka, N. Kaio, and S. Osaki. *Statistical estimation algorithms for repairs-time limit replacement scheduling under earning rate criteria.* Computers & Mathematics with Applications **51**(2), 345–356 (2006). 81

[92] Koichiro Rinsaka and Tadashi Dohi. *Optimizing software rejuvenation schedule based on the kernel density estimation.* Quality Technology & quantitative Management **6**(1), 55–65 (2009). 81

[93] F. Jensen and N.E. Petersen. *Burn-in : an engineering approach to the design and analysis of burn-in procedures.* J. Wiley (1982). 106

[94] J. Mi. *Burn-in and maintenance policies.* Advances in Applied Probability **26**(1), 207–221 (1994). 106

[95] J.H. Cha. *On a better burn-in procedure.* Journal of applied probability **37**(4), 1099–1103 (2000). 106

[96] J.H. Cha. *On optimal burn-in procedures - a generalized model.* IEEE Transactions on Reliability **54**(2), 198–206 (2005). 106

[97] I.B. Sidibe, Khatab A., and K.H. Adjallah. *Kernel based estimation method for age replacement policy.* 5th International Conference on Industrial Engineering and Systems Management, Rabat, Morocco (october 2013). 109

[98] I.B. Sidibe, Khatab A., and K.H. Adjallah. *Kernel estimation based method for the optimization of the system age preventive replacement policy.* International Journal of Reliability, Quality and Safety Engineering. 109

[99] T. Nakagawa and S. Osaki. *Optimum replacement policies with delay.* Journal of Applied Probability **11**(1), 102–110 (1974). 112, 113

[100] C. Tilquin and R. Cléroux. *The block replacement model with inactivity periods and general cost structure.* The Canadian Journal of Statistics / La Revue Canadienne de Statistique **2**(2), 197–213 (1974). 112, 113

[101] D. Ait-Kadi and R. Cléroux. *Optimal block replacement policies with multiple choice at failure.* Naval Research Logistics (NRL) **35**(1), 99–110 (1988). 113

[102] D. Ait-Kadi and R. Cléroux. *Replacement strategies with mixed corrective actions at failure.* Computers & Operations Research **18**(2), 141–149 (1991). 113

[103] I.B. Sidibe, Khatab A., and K.H. Adjallah. *Reliability and preventive maintenance analysis of weibull distributed lifetime systems.* 11th IFAC Workshop on Intelligent Manufacturing Systems, Sao Paulo, Brasil (May 2013). 124

[104] G. Saporta. *Probabilites, analyse des donnees et statistique.* Ed. Technip (2011). 133

[105] P. Hall and M.P. Wand. *On the minimization of absolute distance in kernel density estimation.* Statistics & Probability Letters **6**(5), 311–314 (1988). 134

[106] L. Devroye. *The double kernel method in density estimation.* Annal. Inst. Henri Poincare **25**, 533–580 (1989). 134

[107] J. Beirlant and L. Devroye. *On the impossibility of estimating densities in the extreme tail.* Statistics & Probability Letters **43**(1), 57–64 (1999). 134

[108] P. Sarda. *Smoothing parameter selection for smooth distribution functions.* Journal of Statistical Planning and Inference **35**(1), 65–75 (1993). 136

[109] N. Altman and C. Léger. *Bandwidth selection for kernel distribution function estimation.* Journal of Statistical Planning and Inference **46**(2), 195–214 (1995). 136

[110] P. Hall and W.R. Schucany. *A local cross-validation algorithm.* Statistics & Probability Letters **8**(2), 109–117 (1989). 138

Résumé

Dans la littérature, plusieurs travaux ont été développés autour de la modélisation, l'analyse et la mise en place de politiques de maintenance pour les équipements sujets à des défaillances aléatoires. Ces travaux occultent souvent les réalités industrielles par des hypothèses telles que la connaissance a priori des distributions paramétriques des durées de vie et l'homogénéité des conditions d'exploitation des équipements. Ces hypothèses sont restrictives et constituent une source de biais parce qu'elles conditionnent l'analyse statistique des politiques de maintenance. Dans ce présent travail de thèse, de telles hypothèses sont relaxées pour permettre la prise en compte et la mise en valeurs des informations dérivant directement des données de durées vie issues de l'exploitation de l'équipement et ce sans passer par un modèle paramétrique intermédiaire. L'objectif de ce travail de thèse consiste alors en le développement de modèles statistiques et d'outils efficaces pour l'analyse des politiques de maintenance basées sur les données de durées de vie hétérogènes. Nous proposons en effet une démarche complète d'analyse de stratégies de maintenance en partant des données de durées de vie jusqu'à l'obtention des politiques optimales de maintenance en passant par une phase d'estimation des lois de probabilité. Les politiques de maintenance considérées sont appliques à des équipements usagés évoluant dans des environnements d'exploitation distingués par leur niveau de sévérité. Dans ce contexte, un modèle mathématique est proposé permettant d'évaluer et d'analyser théoriquement les coûts unitaires d'une stratégie de maintenance particulière dite de type âge. Cette analyse a permis d'établir les conditions nécessaires et suffisantes garantissant un âge optimal de remplacement préventif de l'équipement. Les coûts unitaires de maintenance sont complètement estimés par la méthode du Noyau de Parzen. Cette méthode d'estimation est non-paramétrique et définie par une fonction noyau et un paramètre de lissage. Il est également montré, dans nos travaux de recherche, que cet estimateur garantit une faible propagation des erreurs induites par le paramètre de lissage. Les résultats obtenus par la méthode du Noyau de Parzen sont proches des valeurs théoriques avec un faible coefficient de variation. Des extensions de la première politique de maintenance sont également proposées et étudiées. Ce travail de thèse s'achève par la proposition d'une approche permettant de mesurer et d'analyser le risque induit par le report d'une maintenance préventive. Ce risque est analysé à travers une fonction risque proposée.

Mots clés : données hétérogènes, estimation, méthode du Noyau, politique de maintenance, équipement usagé

Abstract

In the reliability literature, several researches works have been developed to deal with modeling, analysis and implementation of maintenance policies for equipments subject to random failures. The majority of these works are based on common assumptions among which the distribution function of the equipment lifetimes is assumed to be known. Furthermore, the equipment is assumed to experience only one operating environment. Such assumptions are indeed restrictive and may introduce a bias in the statistical analysis of the distribution function of the equipment lifetimes which in turn impacts optimization of maintenance policies. In the present research work, these two particular assumptions are relaxed. This relaxation allows to take into account of information related to conditions where the systems is being operating and to focus on the statistical analysis of maintenance policies without using an intermediate parametric lifetimes distribution. The objective of this thesis consists then on the development of efficient statistical models and tools for managing the maintenance of equipments whose lifetimes distribution is unknown and defined through the heterogeneous lifetimes data. Indeed, this thesis proposes a framework for maintenance strategies determination, from lifetimes data acquisition toward the computation of optimal maintenance policies. The maintenance policies considered are assumed to be performed on used equipments. These later are conduct to experience their missions within different environments each of which is characterized by a degree of severity. In this context, a first mathematical model is proposed to evaluate costs induced by maintenance strategies. The analysis of these costs helps to establish the necessary and sufficient conditions to ensure the existence of an optimal age to perform the preventive maintenance. The maintenance costs are fully estimated by using the Kernel method. This estimation method is non-parametric and defined by two parameters, namely the kernel function and the smoothing parameter. The variability of maintenance costs estimator is deeply analyzed according to the smoothing parameter of Kernel method. From these analyses, it is shown that Kernel estimator method ensures a weak propagation of the errors due to the computation of smoothing parameter. In addition, several simulations are made to estimate the optimal replacement age. These simulations figure out that the numerical results from the Kernel method are close to the theoretical values with a weak coefficient of variation. Two probabilistic extensions of the first mathematical model are proposed and theoretically discussed. To deal with the problem of delayed preventive maintenance, an approach is proposed and discussed. The proposed approach allows evaluating the risk that could induce the delay taken to perform a preventive maintenance at the required optimal date. This approach is based on risk analysis conduct on the basis of a proposed risk function.

Keywords : heterogeneous lifetimes data, estimation, Kernel method, maintenance policy, used equipment